Design and Synthesis of a Biologically Active Antibody Mimic Based on an Antibody–Antigen Crystal Structure

M. L. Smythe[1] and M. von Itzstein*,[2]

Contribution from the Department of Pharmaceutical Chemistry, Victorian College of Pharmacy, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia

Received July 6, 1993®

Abstract: We have used the crystal structure of an N9 sialidase (antigen)–NC41 (antibody) complex to design a low molecular weight compound that mimics the binding function of the macromolecular antibody. The components of recognition between the antibody and the protein antigen have been analyzed from the energy-refined crystal complex. From this analysis, four amino acid residues on the antibody binding surface, which make direct contact with the active-site loop 368–370 of the antigen, have been identified as contributing the majority of the binding energy of the protein. The designed target compound, a constrained cyclic peptide, which mimics the receptor-bound conformation of these amino acids, has been synthesized and found to inhibit N9 sialidase activity, with a K_i of 1×10^{-4} M.

Introduction

Protein antigen–antibody interaction sites are large, burying approximately 1500 Å^2 of surface area upon complexation.[4] Their combining sites comprise approximately 15–20 amino acid residues for the three-dimensional structures determined to date.[4–8] Clearly, if all of these residues on the protein binding sites are making equal contributions toward the binding energy between the two proteins, then the task of mimicking the protein binding regions with low molecular weight compounds would be difficult. However, results from competitive experiments of monoclonal and polyclonal antibodies suggest that four, or even less, residues are sufficient to define an antigenic epitope.[9] Similarly, synthetic peptides derived from amino acid sequences of complementarity-determining regions (CDR's) may have analogous binding properties to those of the intact antibody.[10] Further, Novotny *et al.* have concluded, from theoretical calculations, that "protein antigenicity involves active, attractive contributions mediated by a few energetic amino acids and a passive surface complementarity contributed by the surrounding surface area".[11]

These data suggest that it may be possible to mimic protein sites, with low molecular weight compounds. The initial study requires the identification of the key ("energetic") amino acid residues on the protein binding sites, followed by the design and synthesis of compounds that mimic the receptor-bound conformation of these key amino acid residues. One would imagine that the protein mimetic would have a lower affinity than the parent protein simply due to the difference in bound surface area (and hence a potential loss of enthalpic and hydrophobic stabilization) of the mimetic versus that of the parent protein.

Our recent interest in sialidases,[12] in particular the design and synthesis of novel influenza sialidase inhibitors as potential anti-influenza drugs,[13] has also led us to the study of an N9 sialidase–NC41 antibody complex. The crystal structure of the complex between influenza viral enzyme sialidase (subtype N9, isolated from an avian source) and the Fab of a monoclonal antibody NC41 (Figure 1) has been determined by Colman *et al.*[7,8] The interface between N9 and NC41 shows a high degree of steric complementarity. There is a groove between antibody NC41 V_L and V_H domains that accommodates a large ridge on the antigenic surface of N9 sialidase. Adjacent to this, CDR H3 of antibody NC41 forms a ridge which matches a pocket on N9 sialidase which is formed by residues 368–370 and 400–403. The enzymic activity of N9 sialidase is inhibited by the binding of antibody NC41,[7,14] and a binding affinity for the complex of 1.2×10^{-7} M has been determined.[15]

Some indication of the mechanism of inhibition by antibody NC41 can be obtained from the crystal structure of the sialidase–sialic acid complex.[16] On the basis of this structure, it has been

* Abstract published in *Advance ACS Abstracts,* February 15, 1994.

(1) Currently a postdoctoral fellow at the Centre for Molecular Design, Washington University, St. Louis, MO, 63130-4899.

(2) To whom correspondence should be addressed.

(3) Abbreviations: Boc, N^{α}-*tert*-butyloxycarbonyl; BOP, benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate; CDR, complementarity-determining region; cHx, cyclohexyl; DCC, 1,3-dicyclohexylcarbodiimide; DCM, dichloromethane; DIPEA, diisopropylethylamine; DMF, *N,N*-dimethylformamide; d_6-DMSO, dimethyl sulfoxide-d_6; DQFCOSY, double quantum filtered correlated spectroscopy; Fab, antigen binding fragment; Fmoc, 9-fluorenylmethoxycarbonyl; fs, femtosecond; HOBt, 1-hydroxybenzotriazole; HPLC, high-performance liquid chromatography; Hz, Hertz; kcal, kilocalorie; Id, nonpolar atoms; K_i, dissociation constant fo inhibitor; pMBHA, *p*-methylbenzhydrylamine; MES, (2-*N*-morpholino)ethanesulfonic acid; MHz, megahertz; min, minute(s); mL, milliliter(s); mmol, millimole(s), mp, melting point; NMM, *N*-methylmorpholine; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; ps, picosecond; RMS, root mean square; ROESY, rotating frame Overhauser enhancement spectroscopy; *t*Bu, *tert*-butyl; TFA, trifluoroacetic acid; TLC, thin-layer chromatography; TOCSY, total correlated spectroscopy; V_H, variable heavy chain; V_L, variable light domain.

(4) Davies, D. R.; Sheriff, S.; Padlan, E. A. *J. Biol. Chem.* **1988**, *263*, 10541–10544.

(5) Amit, A. G.; Mariuzza, R. A.; Phillips, S. E. V.; Poljak, R. J. *Science* **1986**, *233*, 747–753.

(6) Sheriff, S.; Silverton, E. W.; Padlan, E. A.; Cohen, G. H.; Smith-Gill, S. J.; Finzel, B. C.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8075–8079.

(7) Colman, P. M.; Laver, W. G.; Varghese, J. N.; Baker, A. T.; Tulloch, P. A.; Air, G. M.; Webster, R. G. *Nature* **1987**, *326*, 358–363.

(8) Tulip, W. R.; Varghese, J. N.; Laver, W. G.; Webster, R. G.; Colman, P. M. *J. Mol. Biol.* **1992**, *227*, 122–148.

(9) (a) Kabat, E. A. *Ann. N. Y. Acad. Sci.* **1970**, *169*, 43–54. (b) Schechter, I. *Ann. N. Y. Acad. Sci.* **1971**, *190*, 394–419. (c) Hodges, R. S.; Heaton, R. J.; Parker, J. M. R.; Molday, L.; Molday, R. S. *J. Biol. Chem.* **1988**, *263*, 11768–11775.

(10) (a) Taub, R.; Gould, R. J.; Garsky, V. M.; Ciccarone, T. M.; Hoxie, J.; Friedman, P. A.; Shattil, S. J. *J. Biol. Chem.* **1989**, *264*, 259–265. (b) Williams, W. V.; Guy, H.; Rubin, D.; Robey, F.; Meyers, F.; Kieber-Emmons, T.; Weiner, D.; Greene, M. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6488–6492. (c) Williams, W. V.; Moss, D. A.; Kieber-Emmons, T.; Cohen, J. A.; Myers, J. N.; Weiner, D. B.; Greene, M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5537–5541. (d) Saragovi, H. U.; Fitzpatrick, D.; Raktabutr, A.; Hiroshi, N.; Kakn, M.; Greene, M. I. *Science* **1991**, *253*, 792–795.

(11) Novotny, J.; Bruccoleri, R. E.; Saul, F. A. *Biochemistry* **1989**, *28*, 4735–4746.

(12) (a) Chong, A. K. J.; Pegg, M. S.; von Itzstein, M. *Biochim. Biophys. Acta* **1991**, *1077*, 65–71. (b) Chong, A. K. J.; Pegg, M. S.; von Itzstein, M. *Biochem. Int.* **1991**, *24*, 165–171. (c) Chong, A. K. J.; Pegg, M. S.; Taylor, N. R.; von Itzstein, M. *Eur. J. Biochem.* **1992**, *207*, 65–71.

(13) (a) von Itzstein, M.; Wu, W.-Y.; Van Phan, T.; Danylec, J.; Jin, B. *Chem. Abstr.* **1991**, *117*, 49151y. (b) Holzer, C. T.; von Itzstein, M.; Jin, B.; Pegg, M. S.; Stewart, W. P.; Wu, W.-Y. *Glyconconjugate J.* **1993**, *10*, 40–45. (c) von Itzstein, M., *et al. Nature* **1993**, *363*, 418–423.

0002-7863/94/1516-2725$04.50/0

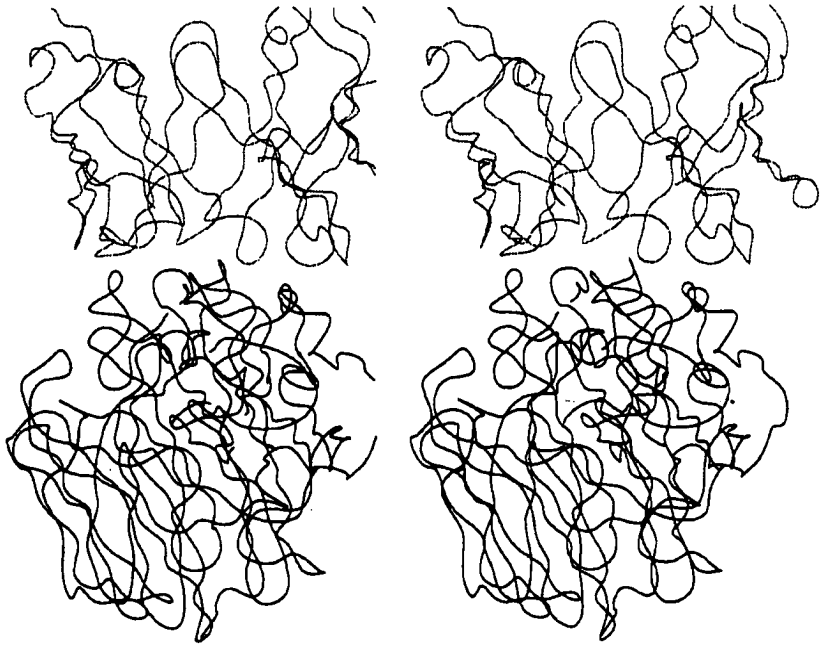


Figure 1. Stereoscopic view of the complex between antigen N9 sialidase and antibody NC41. The upper protein is the variable domains of antibody NC41, and the lower protein is the antigen N9 sialidase.

recently proposed that the antibody NC41 inhibits the enzyme N9 sialidase by binding to the active-site loop 368–370 and so sterically interfering with the approach of the substrate. Analysis of this complex suggests that the approach of the substrate into the active site places the sialoside aglycon moiety near sialidase residues 368–370. Upon binding, the sialoside is known to be conformationally distorted[12] away from a $^2C_5$ conformation, leaving the aglycon positioned directly above the site, well removed from residues 368–370.

We report here the design, synthesis, and biological evaluation of an antibody NC41 mimic based on an analysis of the crystal structure of the N9 sialidase–NC41 antibody complex. It is proposed that this antibody mimic binds to the active-site loop 368–370 of the antigen and inhibits enzyme activity by sterically interfering with the approach of the substrate.

Experimental Section

Analysis of N9 Sialidase–Antibody NC41 Complex. Atomic coordinates of the N9 sialidase–NC41 antibody complex (2.5 Å, *R*-factor 0.191) were kindly provided by Dr. P. Colman and colleagues (Division of Biomolecular Engineering, CSIRO). The complex was energy-refined, using a combination of steepest descents and conjugate gradient minimizations, with the AMBER (v 3.0)[17] software package, to a gradient of less than 0.01 kcal/(mol/Å). The AMBER united-atom force field, a nonbonded cutoff of 8.0 Å, and a distance-dependent dielectric function were employed.

In order to gauge how electrostatically complementary the two protein binding surfaces are in relation to their interactions with solvent, the protein binding sites were solvated with an 8.0-Å layer of equilibrated water molecules using the in-house program CONTACT.[18] These conformations were then minimized to a gradient of less than 0.01 kcal/(mol/Å) (using the AMBER united-atom force field) with the protein held fixed. Molecular dynamic simulations (employing a constant dielectric of 1.0, an 8.0-Å cutoff, and a 1-fs time step) of the solvated protein binding sites were accomplished with the DISCOVER (v 2.6)[19] implementation of the AMBER united-atom force field. The water molecules were subjected to a total of 15 ps of molecular dynamics at 300 K. The protein was held fixed during these simulations.[20] At 1-ps intervals the instantaneous conformation was minimized to a gradient of less than 0.01 kcal/(mol/Å). The lowest energy conformation found was used in the analysis procedure.

The analysis procedure comprised an enthalpic and a nonbonded geometrical calculation of the antibody–antigen interaction in order to determine the key ("energetic") components of the interaction. The process comprised the use of the program CONTACT and the analysis module (anal) of AMBER. The energy-refined complex and the solvated protein conformations were subjected to the analysis process. For the geometrical nonbonded analysis (using program CONTACT), atoms were considered to form intermolecular contacts if their separation distance was within the sum of their van der Waals radii plus 1.0 Å. The program categorized such contacts into charged, polar and nonpolar classes. When any dipole contacts (polar) were considered, the angular component of the contact was calculated.[18] Charged atoms were defined as the charged atoms of residues Glu, Asp, Lys, Arg, and N^{δ} of His. Polar atoms were defined as noncharged heteroatoms and carbons that are bonded to heteroatoms. Nonpolar atoms (Id) were defined as carbon atoms that are solely bonded to carbon atoms. The analysis module of AMBER was used to calculate the enthalpy of interaction of each antibody residue with N9 sialidase. A distance-dependent dielectric and an 8.0-Å cutoff were used for this calculation. Such information allowed us to identify a key loop on the antibody surface that appears to contribute a significant amount of the interaction energy (Table 1).

Design

Organic Scaffold Design. Having identified a key loop (CDR H3) on the antibody surface, the next stage was to design a compound whose conformation was restricted to the receptor-bound conformation of this key loop. The first stage of the design process involved the design of compounds that mimicked the receptor framework (defined in Figure 2) of the key antibody

(14) Colman, P. M.; Air, G. M.; Webster, R. G.; Varghese, J. N.; Baker, A. T.; Lentz, M. R.; Tulloch, P. A.; Laver, W. G. *Immunology Today* **1987**, *8*, 323–326.

(15) Gruen, L. C.; McInerney, T. L.; Jackson, D. C.; Webster, R. G. *J. Protein Chem.* **1993**, *12*, 255–259.

(16) Varghese, J. N.; McKimm-Breschkin, J. L.; Caldwell, J. B.; Kortt, A. A.; Colman, P. M. *Proteins* **1992**, *14*, 327–332.

(17) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Progeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

(18) Smythe, M. L. Design and Synthesis of Protein Mimics. Ph.D. Thesis, Victorian College of Pharmacy, University of Melbourne, Australia, 1992.

(19) Biosym Technologies Inc., 9685 Scranton Rd., San Diego, CA 92121-2777.

(20) It is important to recognize that such an approach used for analysis of protein solvation is based on the assumption that the two proteins would interact in a fashion similar to a "lock and key". In this way the solvated protein conformation would be identical to that found in the complex.

Table 1. Analysis of the Protein Complex[a]

	buried charge–charge	charge–polar	polar–polar	Hbond	charge–Id	polar–Id	Id–Id	enthalpy (kcal/mol)
NC41-N9	1	18	84	26	19	222	49	−63.6[b]
L49-L56	0	0	24	4	0	75	14	−10.5
L91-L96	0	3	10	4	4	21	7	−10.1
H30-H32	0	0	23	3	0	24	8	−7.3
H50-H54	0	0	3	2	0	66	13	−2.8
H96-H99	1	11	23	10	14	61	8	−32.9

[a] A summary of the number of contacts of each loop of the antibody with the antigen. The entire antibody is represented as NC41-N9. Id represent nonpolar atoms. [b] This enthalpy represents the sum of the interaction energy of the individual loops listed in the above table. The total protein–protein interaction energy was calculated as −73.4 kcal/mol.

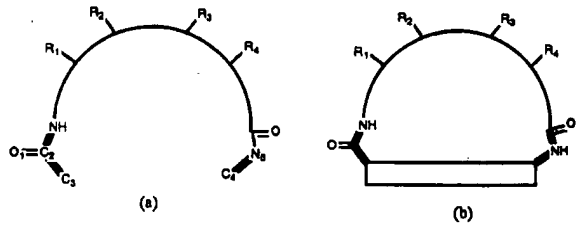


Figure 2. (a) Receptor loop framework (thick lines) describing the initiation and termination of the loops of interest. The receptor framework is defined as the distance between C_2 and N_5 and dihedral angles between the two vectors C2–C3 and C4–N5. For the key antibody loop (CDR H3) the receptor framework was calculated as 5.5 Å (C2–N5 distance), −56.8° (C2–C3–C4–N5 dihedral angle), and −55.6° (O1–C2–C3–C4 dihedral angle). (b) Organic scaffolds are designed that accommodate both the distance spanning the loop (C3–C4) and the C2–C3 and C4–N5 vectors.

loop. Thus, atomic distances and dihedral angles that define the geometry of the receptor framework of the key antibody loop (CDR H3) were calculated. Organic "scaffolds", considered in the design process, were built using the build module in MACROMODEL (v 1.5).[21] The compounds were energy-refined using the implementation of the MM2 force field in MACROMODEL. The dihedral angles and distances defining the receptor framework geometry of the protein loop were then applied as constraints to the molecules. The constrained conformation was minimized to a gradient of less than 0.02 kcal/(mol/Å). A force constant of 5 kcal/(mol/Å) was used for this calculation. These calculations were performed using a distance-dependent dielectric and a cutoff of 13 Å for both van der Waals and electrostatic interactions. The resulting conformation (E_{con}) was then superimposed onto the desired receptor framework, and the root mean square of superimposition over the five receptor template atoms was calculated. If the root mean square was suitable (less than 0.5 Å), then the molecule was subjected to a conformational search. This was necessary in order to gauge the energy accessibility of the receptor framework conformation of the organic scaffold.

Conformational searching was achieved using the multiconformer module of MACROMODEL. The torsion angles defining the rotation about the amine and the carboxyl functional groups were rotated in 10-deg increments. Conformations that had no nonbonded interactions under 1.5 Å were collected. These conformations were energy-minimized to a gradient of 0.02 kcal/(mol/Å) using MM2. If the resulting conformations were unique and were within 5 kcal/mol above the lowest energy conformation found, they were subsequently stored. If the designed organic scaffolds had a suitable root mean square and the E_{con} conformation was within 5 kcal/mol above the lowest energy conformation found (E_{min}), these structures were submitted to the second phase of the design process (template force).

Template Force. After constraining the entering and leaving angles of the loop, by an organic scaffold, the second phase of the design process was to determine how well this scaffold "holds" the peptide in the desired receptor-bound conformation. The constrained cyclic peptides were built using INSIGHT II (v 2.1.0).[19] The receptor conformation of the antibody loop was extracted from the antibody coordinates. The side-chain atoms of the constrained cyclic peptide were template-forced onto the receptor loop side-chain atoms. This involved a constrained minimization of the cyclic peptide to a gradient of less than 0.01 kcal/(mol/Å) using a combination of steepest descents and conjugate gradient minimizations. This was done with the DISCOVER (v 2.6) program. The DISCOVER force field (CVFF, without Morse potentials or crossterms) was used with a template-forcing constant of 20 kcal/(mol/Å). The energy of this conformation (E_{force}) and the root mean square superimposition of the side-chain atoms with the desired receptor conformation were then calculated.

High-Temperature Molecular Dynamics Conformational Search. In order to gauge the energy accessibility of the receptor-bound conformation, and hence how well the scaffold holds the peptide in the desired receptor-bound conformation, a conformational search was conducted. The template-forced conformation (E_{force}) was minimized (with no constraints in place) to a gradient of 0.01 kcal/(mol/Å) prior to molecular dynamics. This conformation was equilibrated for 20 ps at 800 K and molecular dynamics resumed for another 100 ps at that temperature. A 1-fs time step was used, and the amide bonds were constrained to a *trans*[22] geometry with a force constant of 5 kcal/(mol/Å). At every 1 ps a structure was collected and minimized to a gradient of less than 0.01 kcal/(mol/Å), using a combination of steepest descents and conjugate gradient minimizations. The difference between the template-forced conformation (E_{force}) and the lowest energy conformation found (E_{min}) was used to gauge the effectiveness of the scaffold.

Synthesis

Materials and Methods. *p*-Methylbenzhydrylamine (pMBHA) resin (0.82 mmol/g) was used as a polymer support. N^{α}-*tert*-butoxycarbonyl- (Boc) and N^{α}-fluorenylmethyloxycarbonyl- (Fmoc) protected amino acids were purchased from Auspep Pty Ltd., Australia. The Fmoc–Asp(OcHx)–OH and Fmoc–Glu(OcHx)–OH amino acids were prepared from their Boc derivatives by trifluoracetic acid deprotection and Fmoc protection.[23] Fmoc–Asp(OH)O*t*Bu was purchased from Bachem. All reagents and solvents were AR grade and were used without further purification. Mass spectra (FAB) were recorded on a Jeol JMS-Dx 300 double-focusing instrument, with a Jeol-FABMS source and argon as incident particle, in a thioglycerol/glycerol matrix. All infrared spectra were recorded using a Hitachi 370–30 spectrophotometer. Melting points were determined on a Reichter micro-melting point apparatus and are uncorrected. All proton NMR spectra

(21) Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Henderickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440–467.

(22) This was used to eliminate *cis–trans* interconversion, which has been shown to be a useful strategy in high-temperature molecular dynamics. *Discover User Guide Part I*; Biosym Technologies: San Diego, CA, 1992; Version 2.6.

(23) Milton, R. C.; Becker, E.; Milton, S. C. F.; Baxter, J. E. J.; Elsworth, J. E. J. *Int. J. Pept. Protein Res.* **1987**, *30*, 431–432.

were recorded at 300.133 MHz on a Brüker AM-300 or AMX-300 spectrometer. Chemical shifts are quoted as δ values (ppm) relative to an internal tetramethylsilane standard. Coupling constants (J) are given in Hertz, with observed single multiplicity designated as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and b (broad).

Preparation of the Scaffold (3-(((9-Fluorenylmethoxycarbonyl)amino)methyl)benzoic acid) (3) (Fmoc-Amb-OH)

3-(Phthalimidomethyl)benzoic Acid (1).[24] A cold (0 °C) solution of benzoic acid (1.2 g, 10 mmol) and *N*-(hydroxymethyl)-phthalimide[25] (1.77 g, 10 mmol) in concentrated sulfuric acid was reacted overnight at 0 °C. The reaction was quenched by addition of the mixture to ice/water, and the title compound (**1**) (2.3 g, 80%) was collected by filtration and dried *in vacuo*: mp 227–230 °C, lit. mp[24] 228.5–230.5 °C; ^{1}H NMR (d_6-DMSO) 4.83 (s, CH_2), 7.4–8.1 (m, aromatics); IR (KBr) 2964, 1768, 1710, 1452, 1418, 1324, 704 cm^{-1}; MS (FAB) 282 (m/z + 1), 264, 185, 110.

3-(Aminomethyl)benzoic Acid (2).[24] To a stirred solution of 3-(phthalimidomethyl)benzoic acid (**1**) (0.44 g, 1.6 mmol) in a mixture of 2-propanol (14 mL) and water (2.4 mL) was added $NaBH_4$ (0.3 g, 8 mmol). After the reaction was stirred for 24 h, TLC indicated the complete disappearance of starting material. Glacial acetic acid (1.6 mL) was carefully added, and when the exothermic reaction subsided, the flask was stoppered and heated at 80 °C for 2 h.

The crude reaction mixture was loaded onto a Dowex 50 (H^+) column washed with water (750 mL) and then eluted with 1 M NH_4OH (500 mL). Ninhydrin-active fractions were collected and pooled for freeze drying, affording 3-(aminomethyl)benzoic acid (**2**) (0.17 g, 45%), mp 244–246 °C, lit. mp[24] 246–249 °C.

3-(((9-Fluorenylmethoxycarbonyl)amino)methyl)benzoic Acid (3) (Fmoc-Amb-OH). 3-(Aminomethyl)benzoic acid (**2**) (0.75 g, 5 mmol) was dissolved in water (5 mL), and the pH of this solution was adjusted to 8 by the addition of triethylamine. To this solution at 0 °C was added (9-fluorenylmethyl)succinimidyl carbonate (1.5 g, 4.5 mmol) in acetonitrile (5 mL). The pH of the solution was monitored and kept at ≈8, by addition of triethylamine. With the pH constant, the reaction mixture was allowed to stir at room temperature for 3 h. The reaction mixture was filtered, and the solvent was removed under reduced pressure. The isolated residue was added to a vigorously stirred 1 M HCl solution. The resulting precipitate was collected by filtration, dissolved in ethyl acetate, and washed with several portions of 1 M HCl. After removal of solvent under reduced pressure, the product was purified by recrystallization from ethyl acetate/hexane to give 3-(((9-fluorenylmethoxycarbonyl)amino)methyl)-benzoic acid (**3**) (1.27 g, 68%); mp 200–201 °C; ^{1}H NMR (MeOH-d_4) 4.12 (t, 1H, J = 6.8 Hz), 4.24 (s, CH_2–N), 4.28 (d, 2H, J = 6.8 Hz), 7.3 (m, 8H), 7.56 (d, 2H, J = 7.3 Hz), 7.68 (d, 2H, J = 7.44 Hz), 7.82 (dt, H6, J = 7.5, 1.5 Hz), 7.89 (bs, H2); MS (FAB) 374 (m/z + 1), 192, 178; HRMS (FAB) Obsd 373.1292 ($C_{23}H_{19}N_1O_4$), Reqd 373.1315.

Peptide Synthesis. All peptides were prepared manually using a solid-phase peptide synthesis vessel, equipped with a Griffin shaker system.

Synthesis of Antibody Mimic (5). pMBHA resin (5 g, 0.82 mmol substitution) was neutralized by two successive 10-min washes with 10% DIPEA/DCM. The first amino acid (Fmoc-Asp(OH)-O*t*Bu) was attached to the resin, as its preformed HOBt ester, to give a resin substitution of 0.5 mmol/g.[26] The substitution on the resin was calculated using a spectrophotometric technique.[27] The remaining amino groups were blocked by acetylation using acetic anhydride (1.5 mL) in DCM-containing pyridine (3.5 mL) for 20 min at room temperature. On a 2-g (1 mmol) scale of resin, the remaining amino acid residues (including the constraint) were added to the growing peptide chain by the sequential stepwise addition of a 2 molar excess of Fmoc-Asp(OcHx)-OH, Fmoc-Glu(OcHx)-OH, Fmoc-Amb-OH followed by Boc-Phe-OH using standard solid-phase synthetic methodologies.[28] Each coupling reaction was carried out using a twofold excess of BOP and HOBt and a threefold excess of NMM. Following coupling of the last amino acid, the *t*Bu and Boc protecting groups were removed by treatment of the peptide resin with 60% TFA/DCM for 45 min. The peptide resin was washed with DCM (4 × 1 min), DMF (2 × 1 min), 10% DIPEA/DMF (2 × 2 min), and DMF (3 × 1 min). Cyclization was achieved by suspension of the resin in DMF (15 mL) and the addition of a threefold excess of BOP/HOBt and 1.6 equiv of DIPEA. The cyclization reaction was monitored by the quantitative ninhydrin assay and was found to be complete after 5 days. A fresh excess of coupling reagents was added every 24 h. The peptide was cleaved from the resin using anhydrous HF in the presence of anisole as the scavenger. The peptide was extracted with 0.1% TFA–30% acetonitrile/water and subsequently water and then lyophilized to yield 360 mg of crude product. The crude product was purified by reverse-phase HPLC on a μ-Bondpack C18 (10 μm, 125 Å) (25 mm × 100 mm) column, eluted with a linear acetonitrile gradient (20–40%), over 50 min with a flow rate of 4.8 mL/min, with a constant concentration of trifluoroacetic acid (0.1% v/v). The linear gradient was generated with a Waters 600 E system controller. The separation was monitored at 230 nm with a Waters 484 tunable absorbance detector and integration achieved with the Baseline software package. Two major products were isolated and were characterized as the target constrained cyclic peptide (**5**) in 5% yield and an imide byproduct (**6**). The amino acid analysis of the constrained cyclic peptide agreed with expectations. Fraction 1: MS 639 (m/z + 1), corresponding to the target cyclic constrained peptide (**5**) ($C_{30}H_{34}N_6O_{10}$). ^{1}H NMR data (d_6-DMSO) are summarized in Table 3. Fraction 2: MS 621 (m/z + 1), corresponding to the imide byproduct (**6**) ($C_{30}H_{32}N_6O_9$); ^{1}H NMR (d_6-DMSO) 2.0 (m, Glu C$_\beta$H$_2$), 2.40 (t, Glu, C$_\gamma$H$_2$), 2.54 (m, Asp C$_\beta$H″), 2.54 (m, Asn C$_\beta$H″), 3.06 (m, Phe C$_\beta$H″), 3.06 (m, Asp C$_\beta$H′), 3.06 (m, Asn C$_\beta$H′), 3.35 (dd, Phe C$_\beta$H′), 4.05 (m, Amb–CH$_2$′), 4.30 (m, Amb–CH$_2$′), 4.37 (m, Asp C$_\alpha$H), 4.59 (m, Glu C$_\alpha$H), 4.75 (m, Amb–CH$_2$″), 4.75 (m, Phe C$_\alpha$H), 5.24 (m, Asn C$_\alpha$H), 7.18–7.55 (m, aromatic), 7.55 (d, Phe NH), 7.90 (d, Glu NH), 8.53 (m, Amb NH), 9.42 (m, Asp NH).

Synthesis of Linear Peptide Ac-Glu-Asp-Asn-Phe-NH$_2$ (7). pMBHA resin (2 g, 0.46 mmol substitution) was neutralized by two successive 10-min washes with 10% DIPEA/DCM. The first amino (Boc-Phe-OH, 0.49 g, 1.84 mmol) was attached to the resin as its symmetrical anhydride. The degree of coupling on the resin was determined by quantitative ninhydrin assay, and the resin was blocked by acetylation with acetic anhydride. On a 2-g (0.92 mmol) scale of resin, the remaining amino acid residues were added using standard solid-phase methodologies.[28] Boc-Asn-OH was coupled as its HOBt ester, and the remaining amino acids (Boc-Asp(OBzl)-OH and Boc-Glu(OBzl)-OH) were coupled as symmetrical anhydrides. After addition of the last amino acid, it was deprotected by treatment with 40% TFA/DCM, neutralized with a 10% DIPEA/DCM wash, and acety-

(24) Oda, R.; Teramura, K.; Tanimoto, S.; Motaki, N.; Suda, H.; Matsuda, K. *Bull. Inst. Chem. Res., Kyoto Univ.* **1955**, *33*, 117–122; *Chem. Abstr.* **1957**, *51*, 11355d.

(25) Buc, S. R. *J. Am. Chem. Soc.* **1947**, *69*, 254–256.

(26) Effects of resin substitution on peptide yields: Plaue, S. *Int. J. Pept. Protein Res.* **1990**, *35*, 510–517.

(27) Meienhofer, J.; Waki, M.; Hiemer, E. P.; Lambros, T. J.; Makofske, R. C.; Chang, C.-D. *Int. J. Pept. Protein Res.* **1979**, *23*, 35–42.

(28) (a) Bodanszky, M. *Principles of Peptide Synthesis*; Springer-Verlag: Heidelberg, Germany, 1984. (b) Barany, G.; Merrifield, R. B. In *The Peptides*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1980; Vol. 2. (c) Stewart, J. M.; Young, J. D. *Solid State Phase Peptide Synthesis*, 2nd ed.; Pierce Chemical Company: Rockford, IL, 1984. (d) Barany, G.; Kneib-Cordonier, N.; Mullen, D. G. *Int. J. Pept. Protein Res.* **1987**, *30*, 705–739.

lated with acetic anhydride. The peptide was cleaved from the resin using anhydrous HF in the presence of the scavenger, anisole. The peptide was extracted with 0.1% TFA–30% acetonitrile/water and subsequently water and then lyophilized to yield 260 mg of crude product. The crude product was purified by reverse-phase HPLC, using a linear acetonitrile gradient (10–50%) over 60 min with a flow rate of 4.8 mL/min to give **7** (19%). Amino acid analyses were consistent with the desired product: MS 565 ($m/z + 1$) $C_{24}H_{32}O_{10}N_6$; ^{1}H NMR (d_6-DMSO) δ 1.71 (m, Glu $C_\beta H_2$), 1.85 (s, CH_3), 2.24 (t, Glu $C_\gamma H_2$), 2.33 (dd, Asn $C_\beta H'$), 2.49 (m, Asn $C_\beta H''$), 2.49 (m, Asp $C_\beta H'$), 2.62 (dd, Asp $C_\beta H''$), 2.79 (dd, Phe $C_\beta H''$), 3.08 (dd, Phe C_β'), 4.20 (m, Glu $C_\alpha H$), 4.28 (m, Phe $C_\alpha H$), 4.41 (m, Asn $C_\alpha H$), 4.48 (m, Asp $C_\alpha H$), 6.92 (s, amide NH), 7.21 (s, amide NH), 7.91 (d, Phe NH), 7.97 (d, Asn NH), 8.05 (d, Glu NH), 8.22 (d, Asp NH).

Characterization and Structure Determination of Antibody Mimic. A 16 mM sample of the antibody mimic (**5**) was prepared by dissolving 7 mg of compound **5** in 0.7 mL of d_6-DMSO. ^{1}H NMR spectra were recorded with a Brüker AMX-300 spectrometer and processed on a Silicon Graphics 4D/25 workstation using the software package FELIX.[29] These experiments included a TOCSY spectrum with a mixing time of 80 ms, a DQFCOSY, and three ROESY spectra with mixing times of 250, 150, and 80 ms. Apart from the temperature studies, all experiments were recorded at 30 °C. All spectra were run with a spectral width of 3030.3 Hz. 1D spectra were recorded with 16K data points, multiplied by a 60° phase-shifted sine bell squared function, and zero-filled to 32K prior to Fourier transformation. The DQFCOSY of the antibody mimic (**5**) was acquired with 4096 data points for each of the 800 t_1 values and 64 transients. Prior to Fourier transformation the time domain was multiplied in both dimensions by a 90° phase-shifted sine bell. The length of the window function in all 2D experiments was adjusted to reach zero at the last experimental point in the t_1 and t_2 directions. $^3J_{NH-CH}$ coupling constants were measured for the antibody mimic (**5**) by selecting appropriate traces from the DQFCOSY spectrum and subjecting each slice to an inverse Fourier transformation, zero-filling to 8K, applying a 60° phase-shifted sine bell, and finally a Fourier transformation. The $^3J_{NH-CH}$ were then measured directly from the modified traces (Table 3). The TOCSY and ROESY experiments of the antibody mimic (**5**) were acquired with 1024 data points for each of the 400 t_1 values and 32 transients. The temperature dependence of the amide protons were measured at 25, 30, 37, and 42 °C.

Solution Conformation. A total of 15 approximate interproton distance constraints, including 13 interresidue constraints were derived from the ROESY spectra.[18] These distances were classified as 2.7, 3.5, and 5.0 Å for strong, medium, and weak crosspeaks.[30] Pseudoatoms (average atoms) were used for the C_β protons of Asp, Asn, and Phe, and the C_γ protons of Glu, and the CH_2 of Amb. With the constraints in place the conformation was minimized, using steepest descents and conjugate gradients, down to a gradient of less than 0.01 kcal/(mol/Å) using the DISCOVER program. A distance-dependent dielectric, a non-bonded cutoff of 12 Å, and a harmonic force constant of 100 kcal/(mol/Å) were used for these calculations. This conformation was equilibrated for 20 ps of restrained molecular dynamics at 800 K and resumed for 100 ps with a 1-fs time step at this temperature. At every 1 ps the instantaneous structure was minimized to a gradient of less than 0.01 kcal/(mol/Å) with the penalty functions in place. The lowest energy conformation obtained was minimized to a gradient of less than 0.01 kcal/(mol/Å) without any constraints in place. The resulting conformation was found to be consistent with the NMR-derived distance constraints.

(29) *Felix 1.0 program*; Hare Research, Inc.: Woodinville, WA, 1990.
(30) Kuntz, I. D.; Thomason, J. F.; Oshiro, C. M. *Methods Enzymol.* **1989**, *177*, 199–205.

Biological Data. Influenza virus N9 sialidase was incubated for 30 min at 37 °C in a buffer (MES, 6 mM $CaCl_2$, pH 6.5) containing either 30 mg/mL or 70 mg/mL fetuin and varying concentrations of inhibitor (antibody mimic (**5**) and linear peptide (**7**)). The sialic acid released by the enzyme was assayed according to the method developed by Warren[31] and then subsequently modified by Aminoff[32] and Aymard-Henry *et al.*[33]

Results and Discussion

Analysis. The determination of the components of recognition of the N9 sialidase–NC41 antibody complex involved a geometrical nonbonded analysis and an enthalpic calculation. The data from this analysis are summarized in Table 1. These data show, that in comparison to the antibody, CDR H3 (H96–H99) comprises 100% of the buried charge–charge contacts, 79% of the charge–polar contacts, 28% of the polar–polar contacts, 78% of the charge–nonpolar contacts, 16% of the nonpolar–nonpolar contacts, and 44% of the hydrogen bonds. Hydrogen bonds are considered as a separate class of polar contacts in this analysis procedure. The intermolecular energy of the interaction was calculated using the anal module of AMBER. This antibody loop comprised 52% of the total intermolecular energy of the entire antibody. Thus, enthalpically at least four amino acids seem to contribute a significant portion of the interaction energy of the complex.

It has been suggested that one of the greatest free energy deficits in protein complexation results from the requirement to desolvate hydrophilic functional groups.[34,35] This is a result of water participating in strong interactions with hydrogen-bond donors and acceptors as well as in neutralizing charged side chains. Due to the geometric constraints of protein molecules, the hydrophilic side chains buried in the protein complex would participate in fewer hydrogen bonds than they would in the undissociated form, when both donors and acceptors are fully water accessible. As a result, the desolvation of these hydrophilic groups, that do not hydrogen bond within the complex, is an expensive component of complexation. A crude estimation of this unfavorable component of complexation can be achieved by identifying the hydrogen-bond donors and acceptors on the protein-combining sites that are not hydrogen bonding within the complex but are hydrogen bonding with water in the undissociated state. This should give some indication of how many hydrogen bonds are lost on complex formation, in comparison to those made with water, and which residues are responsible for this loss.

The hydrogen-bonding capacity of the antibody recognition site in the undissociated state was characterized by solvating the antibody binding site with an 8.0-Å layer of water molecules. The lowest energy conformation found, from the molecular dynamics study, was used in the analysis procedure. It was found that loop CDR H3 only buries two atoms upon complexation that are hydrogen bonding to the solvent and not with the protein, in comparison to the entire antibody which buries 14 such atoms. This loop makes a total of 10 hydrogen bonds to the antigen, and therefore does not appreciably contribute to complex destability.

Some measure for the role of the electrostatic interactions in protein–protein complexation has been obtained by the above analysis procedure. However, it has been suggested by others that hydrogen bonds and electrostatic and van der Waals interactions contribute little to the stability of protein interactions.[34,35] This is because these interactions replace similar ones made with solvent molecules in the unbound species. Their main role is thought to confer specificity. One aspect of this study was

(31) Warren, L. *J. Biol. Chem.* **1959**, *234*, 1971–1975.
(32) Aminoff, D. *Biochem. J.* **1961**, *81*, 384–392.
(33) Aymard-Henry, M.; Coleman, M. T.; Dowdle, W. R.; Laver, W. G.; Schild, G. C.; *Bull. W. H. O.* **1973**, *48*, 199–202.
(34) Chothia, C. H.; Janin, J. *Nature* **1975**, *256*, 705–708.
(35) Janin, J.; Chothia, C. H. *J. Mol. Biol.* **1976**, *100*, 197–211.

to investigate the hydrophobic contributions[34-37] of complex formation. Whilst under a great deal of conjecture lately,[38] one view is that the hydrophobic effect can be considered as an absence of hydrogen bonding between nonpolar molecules and water, rather than a favorable interaction between the nonpolar groups themselves. The hydrophobic effect can be considered as an entropic force.[37] The favorable entropy term arises due to the release of water molecules upon complexation. This is consistent with the "cavity"[37,39] model, in which the first step may be regarded as the formation of a cavity in water into which the solute will fit. The free energy of formation of such a cavity in liquid water will be large because it requires the separation of strongly interacting water molecules. Once the cavity is formed and the solute placed in it, the solvent will undergo any further changes that reduce the free energy of the system. With a polar solute, hydrogen bonds and other electrostatic interactions will compensate for the initial energy required to form the cavity. However, a nonpolar solute will gain only the minor van der Waals forces with the solvent. To compensate for the absence of favorable electrostatic interactions with the solute, the solvent surrounding it will rearrange to form the most extensive number of hydrogen bonds between water molecules. The water molecules are therefore "immobilized" on the binding surface and significantly contribute to the hydrophobic effect when two nonpolar molecules interact.

The preceding analysis of the solvation of the binding surface of antibody NC41 permits a crude estimation of the hydrophobic effect. As an initial oversimplified estimate of hydrophobicity, an estimation of the number of "immobilized" water molecules on the protein binding surface was required. An "immobilized" water molecule is defined as a water molecule that does not hydrogen bond with the protein and interacts with the nonpolar atoms of the protein. It is these water molecules that are thought to contribute a significant portion of the hydrophobic effect.[40]

The previously calculated low-energy solvated conformation was analyzed by the program CONTACT in order to identify the "immobilized" water molecules on the protein binding surface. The antibody binding site comprises 63 such water molecules. Of these, 21 (one-third) are with the loop CDR H3. This simplistic estimation of hydrophobicity tends to suggest that loop CDR H3 contributes approximately one-third of the hydrophobic force of the entire protein. In addition, upon binding to the antigen this loop would displace approximately 17 immobilized waters on the antigenic surface. Upon complexation this loop buries a total of 204 Å^2 of surface area.[8] This is approximately one-third of the entire protein, and such information qualitatively suggests that this loop would contribute a significant amount[34,41,42] to the hydrophobic effect.

Collectively, from this analysis, it appears that one loop on antibody NC41 (CDR H3) contributes a significant portion of the interaction energy. In the complex, CDR H3 forms a ridge which matches a shallow depression on N9 lined by residues in the active site loop 367–372, loop 400–403, and loop 430–435. The side chain of Asn H98 of the antibody protrudes from the surface and interdigitates between the side chains of Thr 401 and Trp 403. A buried salt link exists between Lys 432 of sialidase N9 and Asp H97 of the antibody. The complex has been described in more detail elsewhere.[8]

Table 2. Summary of the Design Data (in kcal/mol)

conformation	energy	RMS[a]	E_{min}	induction energy
E_{con}	7.7	0.2	4.6	3.1
E_{force}	97	0.6	89	8

[a] Root mean square of the E_{con} conformation is with the five atoms defining the receptor framework. Root mean square of the E_{force} conformation is over all the side-chain atoms.

As mentioned above, the antibody loop (CDR H3) makes direct contact with the N9 active-site loop 368–370. The analysis data, taken together with the hypothesis that antibody NC41 inhibits N9 sialidase activity by binding to the active-site loop 368–370,[16] present the opportunity for the design and synthesis of a compound that mimics these antibody loop amino acids. This mimic may also bind to the active-site loop 368–370 and therefore sterically interfere with enzyme catalysis.

Design. In the preceding analysis section we have described a theoretical technique used to identify several amino acids out of the 20 or so which comprise the antibody recognition site that appear to be responsible for molecular recognition. On the basis of the receptor-bound conformation of these amino acids, known from the protein crystal structure, a design strategy toward suitable cyclic constraints that would hold the peptide sequence in the preferred receptor-bound conformation was required.

Conformational constraints can be separated into two classes.[43,44] The first group, "local" conformational constraints, is used to modify the local conformation to a specific or highly restricted conformation.[43] Their utilization is best undertaken within the context of a structure that already provides a measure of conformational stability, such as a β-turn or an α-helix. The second group, "global" constraints,[43,44] promotes or stabilizes more comprehensive structural features such as α-helices, β-sheets, β-turns, and loops, which are important or sometimes essential for the biological activity of the compound.[43]

Due to the precise steric and electronic requirements of the N9 sialidase–NC41 antibody recognition,[45] this stabilization should not interfere with the steric and electronic properties of this loop. The problem with the most commonly used global conformational constraints[43,44] is their inherent flexibility due to the inclusion of several rotatable sp^3 bonds. In order to circumvent this problem, it was decided to design rigid organic compounds ("scaffolds") that could bridge the C and N termini of the loop.

Design of Organic Scaffolds. In this study, the novel design approach involved the use of the receptor template geometry (defined in Figure 2) of the key loop as conformational constraints in the first phase of the design process. The atomic distances and dihedral angles defining the receptor framework of the key antibody loop (CDR H3) were calculated as 5.5 Å, 56.8°, and −55.6°, respectively (Figure 2). This data was used to identify rigid organic "scaffolds" that could energetically adopt the distance spanning the loop (C3–C4, Figure 2) and the C2–C3 and C4–N5 vectors (the receptor template). The results showed that 3-(aminomethyl)benzoic acid had a root mean square of 0.2 Å with the five atoms defining the receptor framework, and this conformation was only 3.1 kcal/mol above the lowest energy conformation found (Table 2, Figure 3). It was thus concluded 3-(aminomethyl)benzoic acid is a suitable rigid organic scaffold.

The ability of the constrained cyclic peptide to adopt the desired receptor conformation was tested using the template force technique.[46] In this technique the root mean square difference

(36) Kauzmann, W. *Adv. Protein Chem.* **1959**, *16*, 1–64.
(37) Creighton, T. E. *Proteins Structure and Molecular Properties*; WH Freeman and Co.: New York, 1984.
(38) (a) Privalov, P. L.; Gill, S. J. *Adv. Protein Chem.* **1988**, *39*, 191–234. (b) Privalov, P. L. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 47–69. (c) Privalov, P. L.; Gill, S. J. *Pure Appl. Chem.* **1989**, *61*, 1097–1104. (d) Murphy, K. P.; Privalov, P. L.; Gill, S. J. *Science* **1990**, *247*, 559–561. (e) Muller, N. *TIBS* **1992**, *17*, 459–463.
(39) Gill, S. J. *J. Phys. Chem.* **1985**, *89*, 3758–3761.
(40) Andrews, P.; Tintelnot, M. *Comprehensive Medicinal Chemistry*; **1990**, *4*, 321–348.
(41) Nemethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1962**, *66*, 1773–1789.
(42) Forsythe, K. H.; Hopfinger, J. A. *Macromolecules* **1973**, *6*, 423–437.
(43) Hruby, V. J.; Al-Obedi, F.; Kazmieski, W. *Biochem. J.* **1990**, *268*, 249–262.
(44) Hruby, V. *J. Life Sci.* **1982**, *31*, 189–199.
(45) Webster, R. G.; Air, G. M.; Metzger, D. W.; Colman, P. M.; Varghese, J. N.; Baker, A. T.; Laver, W. G. *J. Virol.* **1987**, *61*, 2910–2916.
(46) Struthers, R. S.; Hagler, A. T In *Conformationally Directed Drug Design. Peptides and Nucleic Acids as Templates or Targets*; Vida, J. A., Gordon, M., Eds.; American Chemical *Society*: Washington, DC, 1984; pp 239–261.

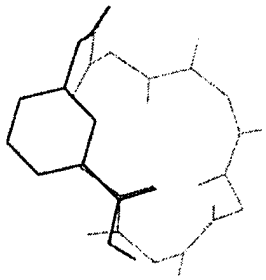


Figure 3. Illustration of superimposition of E_{cons} onto the desired receptor template.

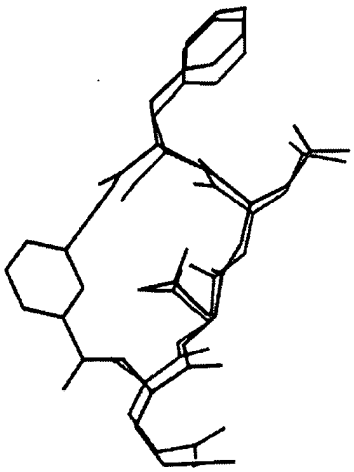


Figure 4. Superimposition of template force conformation (black) onto the desired receptor-bound conformation (gray).

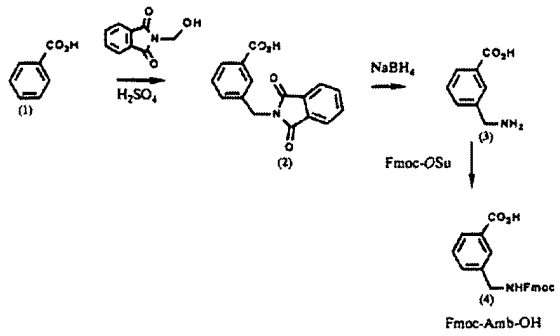
between the analogue (constrained cycle peptide) and the template (desired loop conformation) is minimized simultaneously with the potential energy of the analogue. By minimizing this combined function, the lowest energy conformation (E_{force} conformer) is found that achieves a given fit of the analogue to the template. By conformational searching, and calculation of a low-energy conformer (E_{min} conformer), the energy required (induction energy) to achieve the desired receptor-bound conformation can be calculated. The data from these calculations is summarized in Table 2 and in Figure 4. The amount of energy required to induce the lowest found *in vacuo* conformation into the receptor-bound conformation was calculated as 8 kcal/mol.

Synthesis. The synthesis of the organic scaffold is shown in Figure 5. 3-(Phthalimidomethyl)benzoic acid (**1**) was prepared in 80% yield by the treatment of benzoic acid with *N*-(hydroxymethyl)phthalimide[25,47] in 90% sulfuric acid at 0 °C. The phthaloyl protecting group 3-(phthalimidomethyl)benzoic acid (**1**) was removed under mild conditions by reduction with sodium borohydride.[48] The final step required the protection of the amino functional group as its Fmoc derivative and was achieved by treatment of **3** with (9-fluorenylmethyl)succinimidyl carbonate at a pH of 8.

The syntheses of the precursor peptides for the cyclic lactam analogues of the antibody mimic (**5**) were accomplished by the solid-phase synthetic methods summarized in Figure 5. After neutralization of the pMBHA resin, the β-carboxyl group of the first amino acid (Asp) was coupled to the resin as its preformed HOBt ester. An *N*-capping procedure was then employed, to acylate any free amino groups on the resin. Importantly, it is the β-carboxyl of Asp that is attached to the pMBHA resin; this leaves an α-carboxyl suitably protected for cyclization. Cleavage

(47) (a) Zaugg, H. E. *Synthesis* **1970**, 49–73. (b) Zaugg, H. E. *Synthesis* **1984**, 85–110.

(48) Osby, J.; Martin, M. G.; Ganem, B. *Tetrahedron Lett.* **1984**, *25*, 2093–2096.



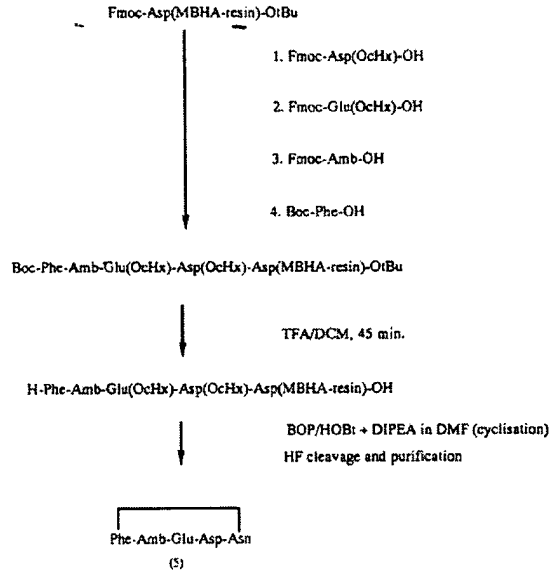


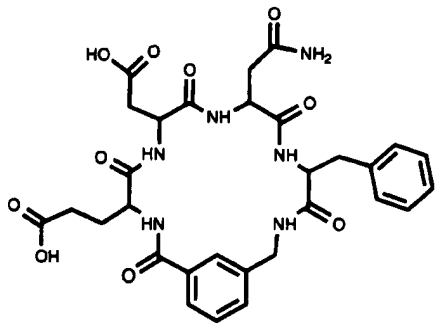
Figure 5. Synthesis of antibody mimic.

from the resin will result in the production of a β-amide and hence an Asn at this position. The remaining amino acids (Fmoc–Glu(OcHx)OH, Fmoc–Amb(OH), and Boc–Phe(OH)) were added using the BOP/HOBt procedure. Prior to cyclization whilst bound to the resin, deprotection of both the *N*-terminal Boc and *C*-terminal *t*Bu functional groups was achieved with 60% TFA/DCM. It was decided to cyclize whilst bound to the resin, as this is reported[26,49] to give products of a higher yield and purity.

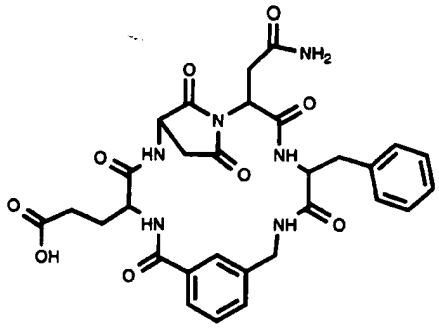
Resin cyclization was initially attempted with DCC/HOBt; however, the major product isolated from this synthesis was the $O \rightarrow N$ acyl migration product.[18,50] Cyclization using BOP/HOBt was then attempted, and the desired constrained cyclic peptide (**5**) was isolated (Figure 6). The cyclization reaction required 5 days to go to completion and was monitored by the quantitative ninhydrin assay. The imide byproduct (**6**) was also detected (Figure 6) and was characterized by the lack of DQFCOSY crosspeak between Asn Cα to the NH region and the ROESY crosspeak from the NH of Asp to the Cα of Glu. Imide formation during cyclization reactions has been previously documented;[26] though it is reported to be significantly reduced by the use of cHx-protected amino acids. The unconstrained peptide (**7**) was synthesized using standard Boc solid-phase peptide synthesis methodologies.

(49) (a) Felix, A. M.; Wang, C. T.; Heimer, E. P.; Fournier, A. *Int. J. Pept. Protein Res.* **1988**, *31*, 231–238. (b) Felix, A. M.; Heimer, E. P.; Wang, C.; Lambros, T. J.; Forunier, A.; Mowles, T. F.; Maines, S.; Campbell, R. M.; Wegrzynski, B. B.; Toome, V.; Fry, D.; Madison, V. S. *Int. J. Pept. Protein Res.* **1988**, *32*, 441–454. (c) Schiller, P. W.; Nguyen, T. M.; Miller, J. *Int. J. Pept. Protein Res.* **1985**, *25*, 171–177.

(50) Burke, T. R.; Knight, M.; Chandrasekhar, B. *Tetrahedron Lett.* **1989**, *30*, 519–522.



(5)



(6)

Figure 6. Structures of the biologically active antibody mimic **5** and the observed imide byproduct **6**.

Table 3. Chemical Shifts[a] and $^3J_{NH-H\alpha}$ Coupling Constants[b] of Antibody Mimic (**5**) in d_6-DMSO at 30 °C

	chemical shift (ppm)			
residue	NH	CαH	CβH	other
Glu	8.60 (6.0 Hz)	4.20	1.97	CγH$_2$, 2.40
Asp	8.44 (11.5 Hz)	4.53	2.64, 2.95	
Asn	7.86 (8.3 Hz)	4.33	2.64	
Phe	8.31 (7.6 Hz)	4.48	2.67, 3.12	C2,3,4,5,6H, 7.18–7.42
Amb	8.48 (5.8 Hz)			CH$_2$', 4.20, CH$_2$'', 4.49, C2H, 7.80, C6H, 7.66, C4H, C5H, 7.42.

[a] Chemical shifts are referenced to the d_6-DMSO peak at 2.49 ppm. [b] Coupling constants were measured from the DQFCOSY spectrum.

Table 4. Temperature Dependence[a] of NH Peaks of Antibody Mimic (**5**) in d_6-DMSO

temp (K)	Glu NH	Amb NH	Asp NH	Phe NH	Asn NH
298	8.544	8.425	8.344	8.230	7.762
303	8.514	8.400	8.336	8.225	7.755
310	8.482	8.373	8.327	8.216	7.745
315	8.450	8.341	8.312	8.206	7.735
coeff (ppm/K)	-6.3×10^{-3}	-5.6×10^{-3}	-2.1×10^{-3}	-1.6×10^{-3}	-1.8×10^{-3}

[a] Chemical shifts (ppm) are referenced to an aromatic peak at 7.42 ppm.

NMR. The ^{1}H spectra of the target molecule (**5**), byproduct (**6**), and linear peptide (**7**) were assigned by a combination of DQFCOSY and ROESY experiments. The assignments of the antibody mimic are summarized in Table 3. The temperature dependencies of the amide protons of the antibody mimic (**5**) were measured in d_6-DMSO and are shown in Table 4. These data indicate that the Asp, Asn, and Phe backbone NH's are strongly hydrogen-bonded, whilst the constraint (Amb) and Glu NH's are in conformational equilibrium between stable hydrogen-bonded conformations and the solvent-exposed environment.[51]

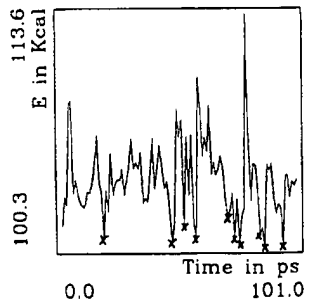


Figure 7. Energy of the 100 minimized structures sampled during the 100-ps (800 K) molecular dynamics search. The low-energy conformations superimposed in Figure 8 are indicated by X's.

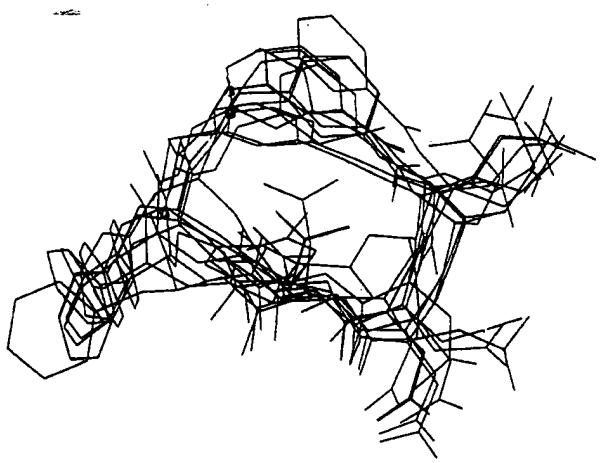


Figure 8. Superimposition of the 10 lowest solution conformations found over a 100-ps molecular dynamics simulation.

The solution conformation of the antibody mimic (**5**) was calculated by a combination of restrained molecular dynamics and restrained molecular mechanics. In total, 100 conformations were collected during the restrained molecular dynamics conformational search. Each of these 100 structures was superimposed onto the side-chain atoms of the desired receptor loop conformation. The root mean square of these 100 conformations, with the desired receptor-bound conformation, had a value between 3.0 and 4.0 Å. A plot of the energy of the 100 structures versus time is presented in Figure 7. The 10 low-energy conformations shown in Figure 7 were obtained for further analysis. The superimposition of these 10 structures is illustrated in Figure 8. The lowest energy conformation found was minimized without the constraints in place. It had an energy of 90 kcal/mol and was found to be consistent with the distance constraints derived from NMR. The energy of the desired receptor-bound conformation was calculated at 95 kcal/mol in the design process. Therefore it would appear as though the organic "scaffold" has stabilized the cyclic peptide to conformations that are energetically close to the desired receptor-bound conformation.

Biological Data. CDR H3 has been shown (theoretically) to be essential for recognition between the antibody and the antigen. Since CDR H3 binds directly to active-site loop (368–370) on the antigen, the antibody mimic may also bind to this region. If this does occur, then it is possible that the antibody mimic can sterically interfere with the approach of the large substrate fetuin and inhibit enzyme activity.

In fact, a K_i of 1×10^{-4} M was determined for the antibody mimic (**5**) using the naturally occurring substrate fetuin against N9 sialidase. These data are shown as Dixon plots in Figure 9.

(51) (a) Hruby, V. J. *Chem. Biochem. Amino Acids, Pept., Proteins* **1974**, *3*, 1–188. (b) Kemp, D. S.; McNamara, P. *Tetrahedron Lett.* **1982**, *23*, 3761–3769.

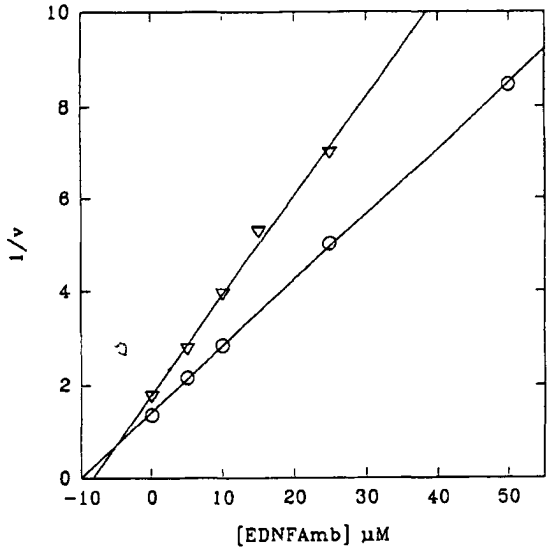


Figure 9. Dixon plot for the inhibition of influenza virus N9 sialidase by the antibody mimic **5** with fetuin as the substrate: ▽, 30 mg/mL fetuin; ○, 70 mg/mL fetuin. $1/v$ is in fluorescence units.

These data clearly indicate that the antibody mimic (**5**) is an inhibitor of the influenza viral enzyme, N9 sialidase. Inhibition produced by the antibody has been previously noted.[7,8,14,15] Since the antibody mimic was designed to bind at the same epitope region as the antibody NC41, it is reasonable to suggest that an inhibition similar to the inhibition observed with NC41 occurs. Indeed the current data supports the notion that the antibody mimic acts in a similar fashion to the antibody NC41. The linear peptide (**7**) did not inhibit enzyme activity, illustrating the importance of the constraint in achieving molecular recognition and hence inhibition of the enzyme.

An alternative explanation to the current interpretation might be that the mimic is binding directly in the active site. However, this is considered unlikely because of the size of the mimic. A study of the dimensions of the active site appears to preclude the mimic from both entering and binding at this site. The antibody mimic was designed to bind at the edge of the active site on loop 368–370 and may inhibit the enzyme by binding to this loop. This is supported by the fact that the mechanism of antibody NC41 inhibiting N9 sialidase is thought to be due to the antibody binding to this same loop.[16]

The mimetic binds approximately 3 orders of magnitude less than that of the parent protein. The antibody mimic contains only four out of the 17 or so residues of the antibody that form contacts with the antigen N9 sialidase. As well, considering the large difference in bound surface area between the antibody–antigen complex and the antibody mimetic–antigen complex, one would imagine that the mimetic would bind with a lower affinity than the parent protein. However, it is possible to improve the affinity of the mimetic by further restricting its available conformational space. The inclusion of local constraints[43,44] may further restrict the conformation of the mimetic to regions about the desired receptor-bound conformation. This would effectively deplete alternative conformations that don't bind and therefore increase binding affinity. This is consistent with the strategy of using global conformational constraints (such as 3-(aminomethyl)-benzoic acid) to stabilize the structural features of the loop, followed by local conformational constraints to further restrict backbone conformations to regions around the desired receptor-bound conformation.

Conclusion

We have analyzed the antibody NC41-N9 sialidase crystal structure and identified one loop on the antibody (CDR H3) that appears to theoretically contribute a significant proportion of the interaction energy between the complex. A constrained cyclic peptide was theoretically designed to mimic the receptor-bound conformation of these key amino acids. The designed organic constraint was synthesized followed by the preparation of the constrained cyclic peptide (antibody mimic) using standard solid-phase synthesis protocols. The desired receptor-bound conformation of the antibody mimic was only 5 kcal/mol higher than the solution (d_6-DMSO) conformation. The biological activity of this low molecular weight compound suggests that it is mimicking the binding function of antibody NC41.

Acknowledgment. We would especially like to thank Prof. P. R. Andrews for his help in the initial stages of the project and the reading of this manuscript. We would also like to thank Dr. P. Colman, Dr. W. Tulip, and Dr. J. Varghese and colleagues for the coordinates of the sialidase N9 antibody NC41 complex and many helpful discussions. We would also like to thank Dr. M. Pegg and Ms. F. Rose for technical assistance in biological testing, Mr. Jeff Dyason for assistance in the preparation of this manuscript, and K. Nielsen for technical assistance in NMR. We would like to thank the National Health and Medical Research Council for financial support, and (M.L.S.) would like to thank Biota Holdings Ltd for a postgraduate fellowship.

Letters in Peptide Science, 4 (1997) 79–84
ESCOM

LIPS 163

Synthesis of α-aspartyl-containing cyclic peptides

Wim D.F. Meutermans, Paul F. Alewood*, Greg T. Bourne, Bernadette Hawkins and Mark L. Smythe

Centre for Drug Design and Development, The University of Queensland, Brisbane, QLD 4072, Australia

Received 13 December 1996
Accepted 13 January 1997

Keywords: Aspartimide formation; β-Peptides; Cyclic peptides; Fluorenylmethyl protection

SUMMARY

α-Aspartyl-containing cyclic pentapeptides were synthesised in high yields using a strategy that maintained fluorenylmethyl protection on the aspartic acid side chain during chain assembly, resin cleavage and cyclisation of the linear precursors. Tetra-*n*-butylammonium fluoride treatment of the fluorenylmethyl-protected cyclic peptides catalysed imide formation, whereas piperidine-induced deprotection resulted in good yields of the target cyclic peptides.

INTRODUCTION

Of the large number of side reactions identified in peptide chemistry, aspartimide formation is one of the most frequently occurring and problematic. It is observed in both Boc- and Fmoc-based syntheses of aspartyl-containing peptides. Imide formation is thought to occur via a nucleophilic attack (Fig. 1) by the backbone nitrogen atom of the i+1 residue on the aspartyl side-chain carbonyl functionality. The reaction may be acid or base catalysed or may proceed spontaneously under neutral conditions in water [1–3]. Subsequent hydrolysis at either of the carbonyl centres of the resulting imide **2** leads to two structural isomers: the α-peptide **3** and the β-peptide **4**. The latter is usually the major product of hydrolysis [4] and may co-elute with the α-peptide on reverse-phase HPLC, making its correct structural identification particularly difficult. Although less common, imide formation can occur in glutamyl-containing peptides, resulting in a mixture of α- and γ-peptides after hydrolysis [5].

The rate of imide formation in a variety of environments is strongly dependent on the nature of the i+1 residue [1,2,6] and of the aspartyl side-chain leaving group (Y) [7–10]. It is fastest when the aspartic acid side chain is activated [11].

*To whom correspondence should be addressed.

Abbreviations: Act, aminocarboxythiophene; Amb, aminomethylbenzoic; *t*-Bu, *tert*-butyl; Boc, *tert*-butoxycarbonyl; Boc_2O, di-*tert*-butyl dicarbonate; BOP, (benzotriazol-1-yloxy)tris(dimethylamino)-phosphonium hexafluorophosphate; DIEA, diisopropylethylamine; DMF, dimethylformamide; Fm, 9-fluorenylmethyl; Fmoc, 9-fluorenylmethyloxycarbonyl; Fmoc-Cl, 9-fluorenylmethyl chloroformate; FmOH, 9-fluorenylmethanol; HBTU, *O*-(1*H*-benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; HF, hydrogen fluoride; HPLC, high-performance liquid chromatography; ISMS, ion spray mass spectrometry; PAM, phenyl acetamidomethyl; TBAF, tetra-*n*-butylammonium fluoride; TFA, trifluoroacetic acid; TIPS, triisopropylsilane.

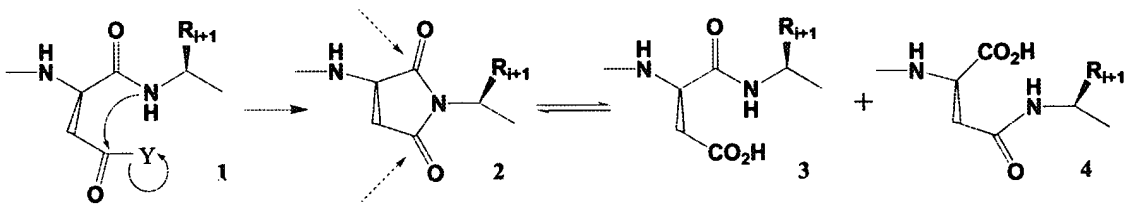


Fig. 1. Aspartimide formation in peptide synthesis. Y = OH, OR or NH_2.

Part of our work on β-turn mimics [12] has necessitated the synthesis of α-aspartyl cyclic peptides which contain constrained amino acids. Two successful strategies have been reported to access α-aspartyl cyclic peptides. Briefly, these include on-resin cyclisation [13] and the solution-phase cyclisation of fully protected linear peptides [14]. Though synthetically elegant, the first method often suffers from low yields in the final cyclisation step while in the second method the fully protected intermediates are often only sparingly soluble in organic solvents or aqueous buffers, thus hindering effective purification and cyclisation.

We considered that these shortcomings could be overcome by solution-phase cyclisation using linear precursors in which only the aspartyl group was protected. In selecting the side-chain protecting group we took the following requirements into consideration: (i) the protected aspartic acid residue should be commercially available; (ii) the protecting group should be stable to strong acid (HF or TFA); and (iii) be readily removed in solution without concomitant imide formation. Here we describe the use of fluorenylmethyl protection of the aspartyl side chain in conjunction with Boc chemistry that allows facile and effective access to α-aspartyl cyclic peptides.

MATERIALS AND METHODS

All resins and Fmoc- and Boc-protected amino acids were purchased from NovaBiochem (Alexandria, NSW, Australia). Boc-Asp(OFm)OH was purchased from Bachem Fein-chemikalien AG (Bubendorf, Switzerland). TBAF, piperidine, *p*-cresol, TFA and TIPS were obtained from Aldrich (Castle Hill, NSW, Australia) and Auspep (Parkville, VIC, Australia). All solvents were of HPLC grade or of equivalent purity. HPLC was carried out on a Waters apparatus with a 600E solvent delivery system and a 484 UV detector. Analytical HPLC was carried out on a Vydac C_{18} reverse-phase column (0.46 × 25 cm) using a 2% linear gradient from 100% buffer A (0.1% TFA in water) to 80% buffer B (90% acetonitrile/10% water, 0.09% TFA) at a flow rate of 1ml/min. Preparative separations on HPLC were carried out on a Vydac C_{18} reverse-phase column (2.2 × 25 cm) using the same linear gradient as described above at a flow rate of 8 ml/min. Mass spectra were recorded on a PE-SCIEX triple quadrupole mass spectrometer using an interface potential of 5.6 kV and an orifice potential of 60–80 V. Samples of approximately 0.1 mg/ml in water/acetonitrile (1:1, 0.1% TFA) were injected directly into the ion spray interface at flow rates of 20–40 μl/min. HPLC fractions were analysed directly without prior treatment. ^{1}H NMR spectra were recorded in aqueous solution (90% H_2O/10% D_2O) on a Bruker ARX 500 spectrometer at 290 K with a peptide concentration of 15 mM. TOCSY spectra were acquired with 4096 complex data points over a spectral width of 8064 Hz, with typically 256–512 increments of 8–32 scans. NOESY spectra were acquired with 4096 complex data points over spectral widths of 6756–8064 Hz with 512–630 increments of 32–64 scans. The data were processed using the Bruker processing package XWIN-NMR.

Synthesis of modified amino acids

Boc-Gly-OH was coupled to 3-amino-4-carbomethoxythiophene [15] with DCC/HOBT in ethyl acetate to give Boc-Gly-Act-OMe in 52% yield. Hydrolysis with LiOH yielded Boc-Gly-Act-OH

(97%): mp 145–150 °C; δ_H ($CDCl_3$, ppm) 1.95 (9H, s, 3 CH_3), 3.70 (2H, m, CH_2), 7.55 (1H, s, NH), 7.95 (1H, d (J = 1 Hz), H_{arom}), 8.27 (1H, d (J = 1 Hz), H_{arom}), 10.85 (1H, s, NH). Boc-protected aminomethyl benzoic acid, Boc-Amb-OH [16], was prepared using Boc_2O via standard procedures [17].

Chain assembly

The linear peptides were synthesised on Boc-Gly-PAM resin (0.5–0.8 mmol/g) by manual stepwise solid-phase peptide synthesis using a fourfold excess of amino acid with HBTU activation and in situ neutralisation [18]. Boc-Gly-Act-OH and Boc-Amb-OH were coupled in a similar fashion (2 equiv amino acid, 2 equiv HBTU (0.5 M in DMF), 3 equiv DIEA, 30 min). Similarly, Fmoc SPPS protocols were used on Fmoc-Gly-Wang resins for the synthesis of linear peptides **6a–d** with Boc-Asp(OFm)-OH as the N-terminal residue. Some sequences were also assembled on Wang resin using Fmoc-protected amino acids and Boc-Asp(OFm)-OH as the N-terminal residue. After TFA cleavage, peptides **6a–d** were separated in good yield and purity.

Cleavage protocols

Peptides synthesised on Wang resin were cleaved by mixing 1 g of dried resin with 90% TFA/10% TIPS (20 ml) for 2–10 h at room temperature. After filtration, TFA was removed under vacuum, ether (15 ml) was added, and the residue was filtered and washed with ether to remove traces of scavenger. The crude residue was then dissolved in a 1:1 mixture of buffers A and B (HPLC), diluted with buffer A and then purified by reverse-phase HPLC. Peptides synthesised on PAM resin were cleaved by HF (18 ml; *p*-cresol (1 ml); *p*-thiocresol (1 ml); –5 °C). After 1 h the HF was removed under vacuum, ether (15 ml) was added, and the residue was filtered and washed with cold ether. The crude peptide was dissolved in a 1:1 mixture of buffers A and B and purified by reverse-phase HPLC.

Cyclisation of Asp(OFm)-protected linear peptides

Linear peptides (0.07 mmol) were dissolved in DMF (70 ml) with stirring at room temperature. DIEA (5 equiv; 61 µl) was then added, followed by BOP (3 equiv; 93 mg). The reaction was left stirring overnight. The DMF was removed under vacuum and the residue was lyophilised from water (0.1% TFA). The crude product was purified by HPLC.

Deprotection of the aspartyl side chain

TBAF The cyclic protected peptides (0.01 mmol) were treated with 0.1 M TBAF solution in DMF (1 ml). After 10 min, water (40 ml) was added, the solution was lyophilised and the products were analysed by HPLC and ISMS.

Piperidine The cyclic protected peptides (0.02 mmol) were treated with 20% piperidine in DMF

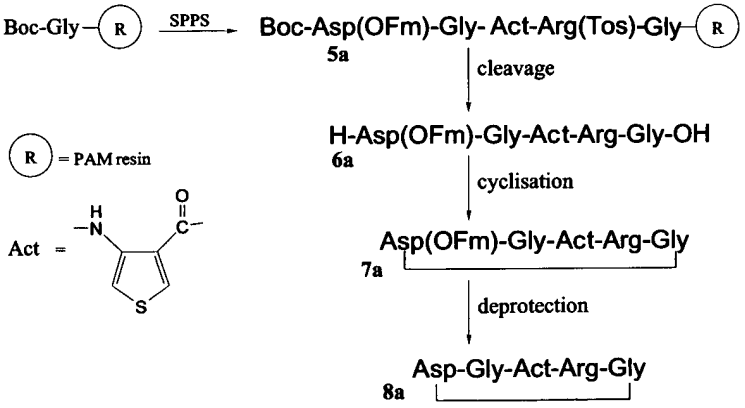


Fig. 2. Synthesis of the cyclic pentapeptide **8a**.

TABLE 1
α-ASPARTYL PEPTIDES AND CYCLISATION YIELDS

	6	7[a,b]
a	H_2N-D(Fm)-G-(Act)-R-G-OH	70 (30)
b	H_2N-D(Fm)-G-(Amb)-R-G-OH	68 (27)
c	H_2N-D(Fm)-G-(β-Ala)-R-G-OH	55 (–)
d	H_2N-D(Fm)-(Amb)-G-R-G-OH	81 (20)
e	H_2N-(Amb)-L-D(Fm)-V-G-OH	49 (4)

[a] Yields are calculated from the weight of isolated material after HPLC purification.
[b] The yields of cyclisation of the aspartyl unprotected linear analogues using the same reaction conditions are in parentheses.

(2 ml). After 1 h of stirring, water (40 ml) was added and the solution was lyophilised and analysed as above.

RESULTS AND DISCUSSION

The synthetic strategy is outlined in Fig. 2 for the synthesis of the cyclic pentapeptide **8a**.

The linear peptides **6a–e** (Table 1) were assembled on PAM resins and, after HF-induced cleavage, isolated in high yields and purity. Glycine was chosen as the C-terminal residue of the linear peptide to avoid racemisation during carboxyl activation in the cyclisation process. Cyclisation of the linear sequences was then achieved at 10^{-3} M in DMF using BOP activation. After HPLC purification, the cyclic Fm-protected peptides **7a–e** were isolated by preparative HPLC in yields ranging from 49 to 81% (Table 1). Conditions for fluorenylmethyl deprotection of **7a** were investigated using either piperidine or TBAF in DMF.

TBAF treatment of peptide **7a** did not lead to the target α-peptide, but resulted mainly in catalysed imide formation accompanied by the loss of fluorenylmethanol, FmOH. The resultant imide **10a** was characterised by NMR and ISMS analysis, showing a loss of 18 mass units from the expected molecular weight (Fig. 3).

Hydrolysis of **10a** led to the formation of two cyclic peptides (**8a**:**9a**) (1:9) of identical molecular weight (an MS/MS examination of both isomers revealed only minor differences in the fragmentation pattern) though with different HPLC retention times (Fig. 5(2A)). A 2D NMR NOESY experiment revealed a strong cross peak from the NH(i+1) to the H^{β} of aspartic acid but not to the H^{α} (Fig. 4), thereby confirming the β-connectivity of the major component **9a**. Thus, TBAF-induced deprotection of **7a** resulted in the preferred formation of the cyclic β-peptide **9a**, with only minor amounts of the target α-peptide **8a** present. A similar observation was made by Kates and Albericio [4], who reported that β-peptides were obtained quantitatively by treating protected linear α-peptides on resin with TBAF in DMF before cleavage.

In contrast, piperidine-induced deprotection of **7a** gave preferentially the desired α-peptide **8a**

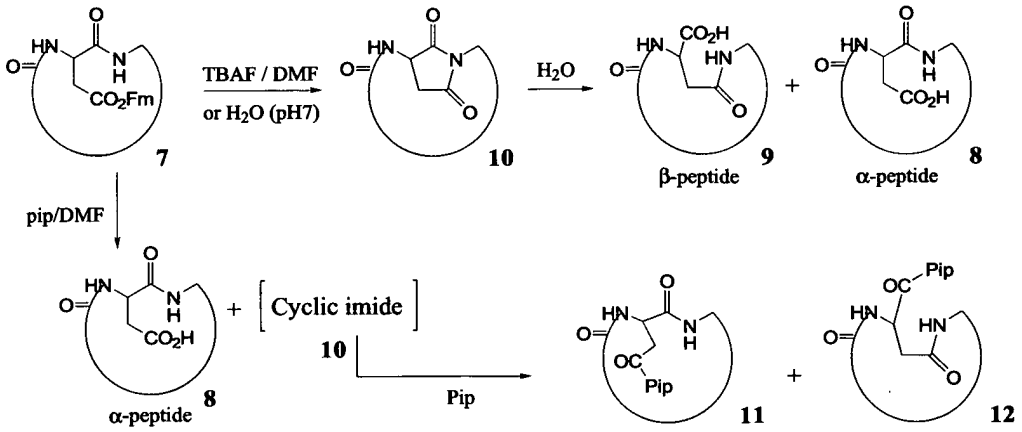
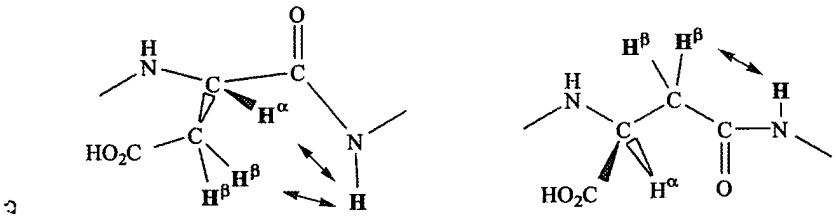


Fig. 3. Deprotection routes for cyclic Fm-protected peptides 7.



8a **9a**

Fig. 4. 2D NMR NOE connectivities for α- and β-connected peptides.

(Fig. 5), with the β-peptide **9a** as a minor product (α:β, 9:1). A 2D NMR NOESY experiment revealed cross peaks from the NH(i+1) to both the H^{α} and the H^{β} of the aspartic acid, supporting the α-connectivity (Fig. 4). The deprotection was further accompanied by the formation of a small amount of imide (through elimination of FmOH) which reacted with piperidine to give the α- and β-adducts **11a** and **12a**. Hydrolysis of the cyclic imide **10a** formed in this piperidine deprotection route, and formation of β-peptide **9a**, is thereby significantly reduced. This result was consistent with the observation that Fm deprotection using 20% piperidine/DMF is rapid ($t_{1/2}$ around 1 min) with respect to base-catalysed imide formation [19,20] (Fig. 5). Similarly, α-aspartyl-containing cyclic peptides **8b–e** were obtained in high yields (75–86%) and purity from piperidine treatment of their Fm-protected precursors. Unlike **7a**, peptides **7b–e** after piperidine treatment gave no detectable amounts of β-peptide in the final product as determined by NMR/HPLC analyses.

The overall yields of the cyclic peptides **8a–e** (31–56%) obtained through this partial protection strategy are significantly higher than the yields obtained from cyclisation of the unprotected analogues (Table 1) or from other recent reports [21].

Stability of cyclic peptides: A caveat

α-Peptide **8a**, β-peptide **9a** and imide **10a** inter-

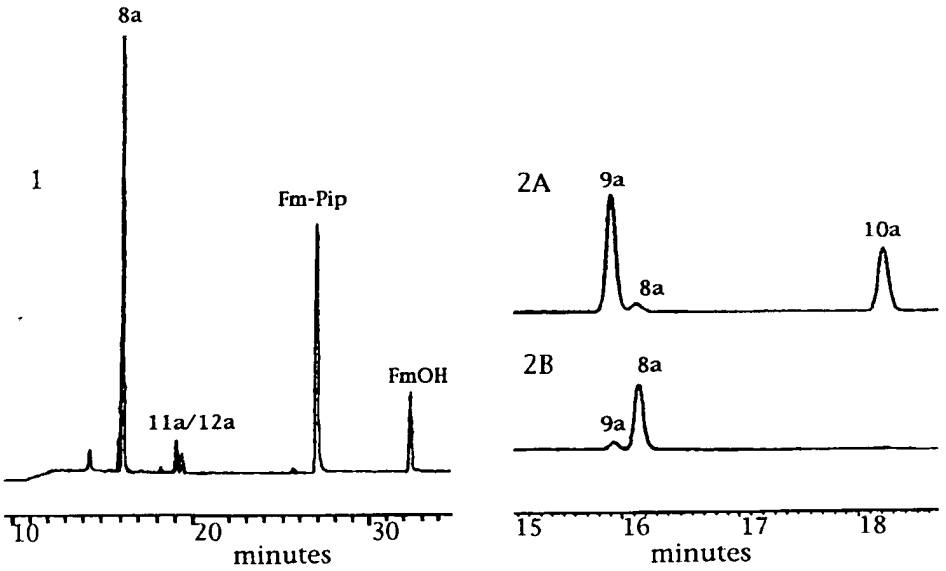


Fig. 5. (1) HPLC profile of the crude reaction mixture from piperidine-mediated deprotection of **7a**; (2) HPLC profile of (A) imide **10a** in water at pH 7 (2 h) and (B) peptides **8a**:**9a** (9:1) obtained from piperidine-mediated deprotection of **7a**.

convert in water at pH < 8. For example, 30% succinimide **10a** is formed after leaving peptides **8a**:**9a** (9:1 ratio) for 7 days at pH 2. Over time this will lead to nearly full conversion of α-peptide **8a** to β-peptide **9a**. Therefore, care needs to be taken when planning experiments involving aqueous solutions of these compounds over extended periods (bioassays or NMR work). In a recent report [22], the effect of pH and buffer on the rate of aspartimide formation of linear tetrapeptides was thoroughly investigated and revealed a bell-shaped curve in the pH–rate profile, with a maximum rate at pH 4.

For the Fm esters **6** and **7** on the other hand, the rate of spontaneous imide formation dramatically increases with increasing pH. The protected cyclic peptide **7a** for instance, when dissolved in water at pH 7, was completely converted in 1 h to the imide **10a**, whereas at pH 2 its half life is around 30 days allowing facile purification and handling without significant loss.

CONCLUSIONS

The fluorenylmethyl protection strategy allows fast and efficient access to α-aspartyl-containing cyclic peptides. Linear peptides containing only Asp(OFm) side-chain protection are readily accessible and sufficiently soluble and stable for effective purification and cyclisation. Piperidine-induced deprotection leads to high yields of the target cyclic products.

ACKNOWLEDGEMENTS

The authors are grateful for the financial support from GlaxoWellcome UK and GlaxoWellcome Australia and wish to acknowledge Dr. Paul Doyle for his continuing interest and for useful discussions.

REFERENCES

1 Tam, J.P., Riemen, M.W. and Merrifield, R.B., Pept. Res., 1 (1988) 6.
2 Bodanszky, M. and Kwei, J.Z., Int. J. Pept. Protein Res., 12 (1978) 69.
3 Geiger, T. and Clarke, S., J. Biol. Chem., 262 (1987) 785.
4 Kates, S.A. and Albericio, F., Lett. Pept. Sci., 1 (1994) 213.
5 Capasso, S., Mazzarella, L., Sica, F. and Zagari, A., J. Chem. Soc., Chem. Commun., (1991) 1667.
6 a. Lauer, J.L., Fields, C.G. and Fields, G.B., Lett. Pept. Sci., 1 (1994) 197.
b. Yang, Y., Sweeney, W.V., Schneider, K., Thornqvist, S., Chait, B.T. and Tam, J.P., Tetrahedron Lett., 35 (1994) 9689.
7 Fujii, N., Nomizu, M., Futaki, S., Otaka, A., Funakoshi, S., Akaji, K., Watanabe, K. and Yajima, H., Chem. Pharm. Bull., 34 (1986) 864.
8 Iguchi, S., Tsuda, Y. and Okoda, Y., Chem. Pharm. Bull., 37 (1989) 2209.
9 Stephenson, R.C. and Clarke, S., J. Biol. Chem., 264 (1989) 6164.
10 Blake, J., Int. J. Pept. Protein Res., 13 (1979) 418.
11 Urge, L. and Otvos Jr., L., Lett. Pept. Sci., 1 (1994) 207.
12 Bourne, G.T., McKie, J.H., Cassidy, P.J., Smythe, M.L., Alewood, P.F. and Andrews, P.R., In Kaumaya, P.T.P. and Hodges, R.S. (Eds.) Peptides: Chemistry and Biology, Mayflower Scientific, Kingswinford, U.K., 1996, pp. 354–355.
13 a. Flanigan, E. and Marshall, G.R., Tetrahedron Lett., 27 (1970) 2403.
b. Osapay, G. and Taylor, J.W., J. Am. Chem. Soc., 112 (1990) 6046.
c. Richter, L.S., Tom, J.Y.K. and Burnier, J.P., Tetrahedron Lett., 35 (1994) 5547.
d. McMurray, J.S., Lewis, C.A. and Obeyesekere, N.U., Pept. Res., 7 (1994) 195.
14 McMurray, J.S., Budde, R.J.A. and Dyckes, D.F., Int. J. Pept. Protein Res., 42 (1993) 209.
15 Baker, B.R., Joseph, J.P., Schaub, R.E., McEvoy, F.J. and Williams, J.H., J. Org. Chem., 18 (1953) 138.
16 Smythe, M.L. and Von Itzstein, M., J. Am. Chem. Soc., 116 (1994) 2725.
17 a. Tarbell, D.S., Yamamoto, Y. and Pope, B.M., Proc. Natl. Acad. Sci. USA, 69 (1972) 730.
b. Carpino, L.A. and Han, G.Y., J. Am. Chem. Soc., 92 (1970) 5748.
18 Schnolzer, M., Alewood, P., Jones, A., Alewood, D. and Kent, S.B.H., Int. J. Pept. Protein Res., 40 (1992) 180.
19 For example, the half life of Fmoc-Val-OH in 20% piperidine/DMF is 6 s: Atherton, E. and Sheppard, R.C., In Udenfriend, S. and Meienhofer, J. (Eds.) The Peptides, Academic Press, London, U.K., 1987.
20 Schon, I., Colombo, R. and Csehi, A., J. Chem. Soc., Chem. Commun., (1983) 505.
21 Jackson, S., DeGrado, W., Dwivedi, A., Parthasarathy, A., Higley, A., Krywko, J., Rockwell, A., Markwalder, J., Wells, G., Wexler, R., Mousa, S. and Harlow, R., J. Am. Chem. Soc., 116 (1994) 3220.
22 Capasso, S., Kirby, A.J., Salvadori, S., Sica, F. and Zagari, A., J. Chem. Soc. Perkin Trans. II, (1995) 437.

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 11138–11142, November 1994
Chemistry

"Libraries from libraries": Chemical transformation of combinatorial libraries to extend the range and repertoire of chemical diversity

(transformation/peralkylated combinatorial libraries/permethylated peptides/peptide library/antimicrobial)

JOHN M. OSTRESH*, GREGORY M. HUSAR*, SYLVIE E. BLONDELLE*, BARBARA DÖRNER*, PATRICIA A. WEBER†, AND RICHARD A. HOUGHTEN*‡

*Torrey Pines Institute for Molecular Studies, San Diego, CA 92121; and †Houghten Pharmaceuticals, San Diego, CA 92121

Communicated by Bruce Merrifield, April 21, 1994 (received for review November 15, 1993)

ABSTRACT The generation of diverse chemical libraries using a "libraries from libraries" concept is described. The central features of the approaches presented are the use of well-established solid-phase synthesis methods for the generation of combinatorial libraries, combined with the chemical transformation of such libraries while they remain attached to the solid support. The chemical libraries that are generated by this process have very different physical, chemical, and biological properties compared to the libraries from which they were derived. A wide range of chemical transformations are possible for peptide-based or other libraries, and an almost unlimited range of useful chemical diversities can be envisioned. In the example presented, the amide functionalities in an existing combinatorial library made up of peptides were permethylated while the library remained attached to the solid-phase support used in its synthesis. After removal of the permethylated mixtures from their solid support, this library, now lacking the typical -CONH- amide bonds of peptides, can be tested in solution with virtually all existing assay systems to identify individual compounds having specific biological activities of interest. An illustration of the use of such libraries is presented, in which the described permethylated library was used to identify individual permethylated compounds having potent antimicrobial activity against Gram-positive bacteria.

Recent innovations in peptide chemistry and molecular biology have enabled libraries consisting of tens to hundreds of millions of peptide sequences to be prepared and used to identify highly active, individual sequences. Such libraries can be divided into three broad categories. (*i*) One category of libraries involves the chemical synthesis of soluble nonsupport-bound peptide and peptoid libraries (1–3). (*ii*) A second category involves the chemical synthesis of support-bound peptide libraries composed of L- or D-amino acid sequences presented on solid supports such as plastic pins (4), resin beads (5), or cotton (6). (*iii*) A third category uses molecular biology approaches to prepare peptides or proteins on the surface of filamentous phage particles or plasmids (7). More recently, the production of small collections of nonpeptidic compounds has been described (8–10).

As first presented by this laboratory, soluble, nonsupport-bound peptide libraries [termed synthetic peptide combinatorial libraries (SPCLs)] appear to be usable in virtually all *in vitro* and even *in vivo* assays. Combinatorial libraries of peptides composed of entirely L-amino acids, entirely D-amino acids, or mixtures of L-, D-, and unnatural amino acids have been developed by using this approach. The successful use of these libraries has been reported for the study of antibody/antigen interactions (1, 2) and for the development of receptor-active opioid peptides (11, 12), enzyme inhibitors (6), and antimicrobial agents (1). The recent development of the soluble positional scanning (PS) approach for the production and screening of SPCLs (2, 12) enables individual, biologically active peptides to be identified in a single screening assay.

In a continuing effort to expand the available repertoire of chemical diversities, we present here an example of a soluble chemical library obtained by the chemical transformation of an existing peptide library. This library, prepared in a positional scanning format, is composed of fully permethylated compounds derived from the direct chemical modification of resin-bound peptide libraries. To our knowledge, a study of individual peptides being permethylated while attached to solid-phase synthesis resins has not yet been reported. The use of chemically transformed libraries is illustrated here by the use of this permethylated PS-SPCL for the identification of potent individual compounds selectively active against the Gram-positive bacteria *Staphylococcus aureus* and *Streptococcus sanguis*.

MATERIALS AND METHODS

The benzyl protecting group was used for the side-chain protection of asparagine, glutamate, serine, and threonine; methoxybenzyl for cysteine; dinitrophenyl for histidine; chlorobenzyloxycarbonyl for lysine; sulfoxide for methionine; tosyl for arginine; formyl for tryptophan; and bromobenzyloxycarbonyl for tyrosine. Other reagents and materials used have been described (13).

Permethylation of Protected Peptides. Twenty model peptide resins (110 mg, 0.1 milliequivalent each) with defined sequences (represented as *O*GGFL-resin, "*O*" individually representing each of the 20 naturally occurring amino acids) were prepared on methylbenzhydrylamine resin (mBHA) (0.90 milliequivalent per g) using *t*-butoxycarbonyl (Boc) chemistry combined with simultaneous multiple peptide-synthesis techniques as described (13). After removal of the final *N*-α-Boc protecting group, each of the side-chain-protected resins were then permethylated as described below. The resin-bound PS-SPCL was prepared as described (2, 12).

In a typical example, AGGFL-mBHA resin (90 mg, 0.1 milliequivalent), contained within an individual polypropylene mesh packet (13), was shaken on a reciprocating shaker (Eberbach, Ann Arbor, MI) at 25°C for 16 hr in a 0.25 M

Abbreviations: PS, positional scanning; RP-HPLC, reversed-phase high-performance liquid chromatography; SPCL, synthetic peptide combinatorial library; MRSA, methicillin-resistant *Staphylococcus aureus*. All amino acids used were of the L configuration unless otherwise noted.

‡To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

solution of NaH/dimethyl sulfoxide (32 ml, 8 milliequivalents). Neat methyl iodide (1.5 ml; 24 milliequivalents) was then added to the reaction mixture, and the methylation reaction allowed to proceed for 15 min at 25°C. After successive washes with dimethylformamide (three times, 5 ml), isopropanol (two times, 5 ml), dichloromethane (three times, 5 ml), and methanol (once, 5 ml), the resin was dried under high vacuum. The resin-bound permethylated peptide was cleaved by using 7.5% (vol/vol) anisole/HF (5 ml) for 1 hr at 0°C, dried under high vacuum, extracted with 10 ml of water, and lyophilized. The permethylated peptide was assayed for purity by reversed-phase high-performance liquid chromatography (RP-HPLC) and identified by laser desorption-mass spectral analysis (Kratos). Individual compounds were purified using preparative RP-HPLC. The PS-SPCL was permethylated, cleaved, and analyzed similarly.

Biological Assays. Individual permethylated compounds and their nonpermethylated counterparts were assayed for their resistance to proteolytic breakdown by trypsin and chymotrypsin. The assays were done in 1 ml of 0.1 M NH_4HCO_3, pH 7.8, at room temperature for 16 hr at a peptide concentration of 1.0 mg/ml. Enzyme-to-peptide concentration was 1:50. The degradation reaction was monitored by RP-HPLC.

The strains *S. aureus* ATCC 29213, methicillin-resistant *S. aureus* (MRSA) ATCC 33591, *S. sanguis* ATCC 10566, *Escherichia coli* ATCC 25922, and *Candida albicans* ATCC 10231 were used in the bioassays. The assays were done in 96-well tissue culture plates (Costar) as described (14).

Hemolytic activity of the individual compounds identified was determined by using a 0.25% suspension of human red blood cells as described (15).

RESULTS

Optimization of Permethylation Conditions. Various methods for permethylation have been described (16, 17). In such permethylations, strongly basic conditions are reported to favor N-methylation over O-methylation. Although unreported in these earlier studies, the synthetic conditions used also permit permethylation of peptides while they remain attached to the solid-phase resins used in their synthesis. The strength of the solid-phase approach (18) is that all excess reagents can be removed by simple wash procedures.

Satisfactory permethylation conditions for resin-bound peptides were studied using AGGFL-NH_2 due to its ease of synthesis, nonreactive side chains, and its availability from other ongoing studies (12). Temperature, reaction time, reagent ratios, and solvents were studied to determine the most effective and mildest conditions for the formation of the amide anions and their subsequent methylation. Temperatures studied ranged from 25°C to 60°C. Reaction times tested for the generation of the amide anions ranged from 20 min to 16 hr. It was found that while complete amide anion formation occurred within 16 hr at 25°C, higher temperatures led to degraded products, as evidenced by mass spectral analysis and RP-HPLC. Methylation reaction times examined ranged from 1 to 30 min. Because permethylation of a free α-amino peptide leads to both methylation of the amide backbone and quaternary salt formation, it was necessary to study the relative rates of these two reactions. It was found that while methylation of the backbone amides was complete within the first minute, longer reaction times were required for the quaternary salt formation. Complete methylation of the backbone amide anions within the first minute was demonstrated with the use of Ac-AGGFL-NH_2. Of the conditions tested, a 16-hr room-temperature treatment of the resin-bound protected peptides, using a 10-fold excess of NaH in dimethyl sulfoxide over the reactive sites of the resin-bound peptide, followed by a 15-min treatment of the resulting amide anions with a 30-fold excess of methyl iodide over the reactive sites, yielded the best results (Fig. 1). Under these conditions AGGFE-NH_2 was obtained in permethylated form in >90% yield and purity.

Permethylation of Model Peptides. Once the permethylation conditions had been selected, 20 model peptide resins (represented as *O*GGFL-NH_2, where "*O*" is one of the 20 naturally occurring L-amino acids) were synthesized. Resin compartmentalization (13) permits the simultaneous permethylation of multiple peptide resins. Mass spectral and RP-HPLC analysis of an HF-cleaved aliquot of the nonpermethylated starting resins indicated that the average crude purity of the nonpermethylated peptides was >95%. The peptide resins were permethylated to determine the stability and susceptibility to modification of the 20 naturally occurring L-amino acids. Mass spectral analyses showed that the nitrogen of each amide bond was methylated, including the C-terminal amide resin linkage. In addition to the quaternization of the α-amino group, small amounts of the mono- and dimethylated α-amino products were also formed, generally to an extent of <15%. While a preliminary exhaustive N-α-methylation of the model peptides before the sodium hydride/methyl iodide treatment eliminated the mono- and dimethylated products, the strongly basic conditions used in the sodium hydride treatment gave partial Hoffman elimination of the quaternary salts formed. In contrast, while exhaustive N-α-methylation after the initial permethylation did not yield the byproducts associated with Hoffman elimination, the quaternization reaction was not driven to completion under the conditions used.

The subsequent cleavage of the permethylated model peptides from the resin yielded the desired products in ~90% purity for amino acids having nonreactive side chains (Fig. 2). Table 1 summarizes the number of methyl groups incorporated (including side-chain modifications) for each of the other analogs. Similar results were obtained using peptide resins in which the amino acid at the second position was varied, demonstrating the lack of a positional dependence on the modification of the side chains (data not shown).

Although the major product of the permethylation of the model peptide containing cysteine was the methyl thioether,

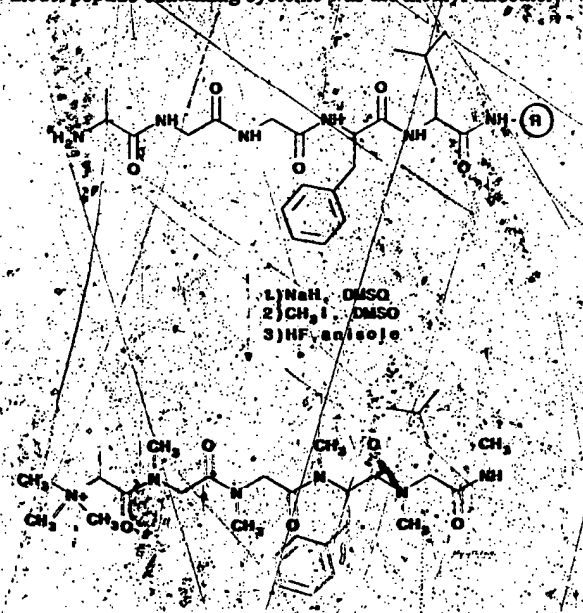


FIG. 1. Illustration of the method used for the permethylation of peptides. DMSO, dimethyl sulfoxide.

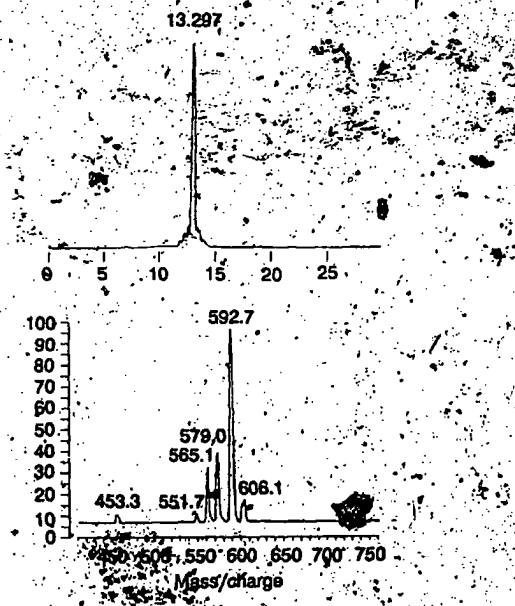


FIG. 2. RP-HPLC and mass spectral analysis of the permethylated form of SGGFL-NH_2.

the presence of unmodified side chain suggests incomplete removal of the methoxybenzyl protecting group. It was found that the benzyl ester side-chain protecting groups of aspartate and glutamate were removed during the sodium hydride treatment. The lower purity found suggests that internal cyclization of these compounds occurred (19). The dinitrophenyl protecting group of histidine was found to be removed by the sodium hydride treatment. Methylation of the histidine side chain gave two major products of identical molecular weight, which are assumed to be the τ- and π-imidazole side-chain methylation products. The low purity of KGGFL-NH_2, along with the presence of low-molecular-weight species as evidenced by mass spectral analysis, indicates additional byproducts are due to internal cleavage by nucleophilic attack of the ϵ-amino group on the peptide backbone.

The sulfoxide of methionine, which is stable to the reaction conditions, also serves to prevent the internal cleavage of the peptides via methyl sulfonium salts of methionine (20). Unexpectedly, some free methionine product was found resulting from reduction of the sulfoxide. The stability of the tosyl group of arginine to the permethylation reaction resulted in three methyl groups being incorporated onto the guanidinium group of the arginine-containing peptide. The tosyl group was then removed during the standard high HF cleavage procedure (21). Methylation of the rings of aromatic amino acids (22) was not seen.

Treatment of the permethylated peptide resins by using standard low HF conditions (23) was found to result in partial cleavage of permethylated peptides from the resin. The lability of secondary amide bonds relative to primary amide bonds under acidic conditions is well documented (24, 26). In addition, the peptide remaining on the resin after low HF treatment was found to be substantially degraded as determined by RP-HPLC and mass spectral analysis after HF cleavage.

Racemization Study. From the literature, it is anticipated that the conditions used here for permethylation would result in minimal racemization or C_α-methylation (17). Because the increased acidity of the C_α-hydrogen of aromatic amino acids such as phenylalanine (17) makes them more prone to racemization and/or C_α-methylation, a test series was devised in which the four possible stereoisomers of GGFL-NH_2 were synthesized and analyzed by RP-HPLC. An aliquot of the GGFL-NH_2 resin was treated with NaH/dimethyl sulfoxide to form the amide anions and then quenched by washing with 1% water/dimethyl sulfoxide. After amino acid cleavage from the resin, the maximum percentage of the D/L, L/D enantiomeric pair, as seen by RP-HPLC, was <0.75%, establishing that the extent of racemization, and therefore potential C_α-methylation, was <1% (Fig. 3).

Enzymatic Susceptibility. The stability of N-permethylated compounds to proteolysis was examined for two permethylated sequences, AGGFL-NH_2 and RGGFL-NH_2; their non-permethylated equivalents were used as controls. Treatment of these four compounds by trypsin and chymotrypsin was monitored by RP-HPLC and mass spectral analysis. Rapid cleavage of AGGFL-NH_2 by chymotrypsin and RGGFL-NH_2 by trypsin was observed (<1 hr), whereas <1% cleavage of the equivalent permethylated peptides was seen after overnight enzyme exposure (3).

Preparation of the Permethylated Library. The six separate positional sublibraries of a resin-bound PS-SPCL (2), each consisting of 20 separate peptide mixtures, were permethylated and cleaved as described. The resulting library consisted of 120 hexamer mixtures, in which one position was defined by each of the 20 permethylated amino acids (represented as *O*), with the remaining five positions made up of mixtures of 18 permethylated amino acids (represented as *X*;

Table 1. Side-chain modifications of amino acids in permethylated peptides

Parent sequence	Parent M_r	Methylated M_r found	Number methyls	Side-chain modification	Nonmethylated purity,† %	Methylated* purity,† %
CGGFL-NH_2	494.8	620.8	9	Methyl thioether	86	50
DGGFL-NH_2	506.7	633.6	9	Methyl ester	97	60
EGGFL-NH_2	520.7	647.8	9	Methyl ester	90	75
HGGFL-NH_2	528.8	655.8	9	Methyl imidazole	98	40
KGGFL-NH_2	519.8	673.0	11	Quaternary salt	99	30
LGGFL-NH_2	504.8	617.2	8	Unmodified	99	81
M[O]GGFL-NH_2	538.8	651.8	8	Unmodified	98	70
NGGFL-NH_2	505.7	646.9	10	Dimethyl amide	85	86
QGGFL-NH_2	519.7	660.0	10	Dimethyl amide	85	80
RGGFL-NH_2	703.3	547.8	11	Trimethyl guanidine	88	75
WGGFL-NH_2	577.8	704.8	9	Methyl indole	98	70
YGGFL-NH_2	554.8	681.5	9	Methyl ether	99	81

*Sum of the mono-, di-, and trimethylated α-amine products.
†Purity of the crude compounds as determined by analytical RP-HPLC. Repetition of these experiments showed a maximum variation of 20% from the values listed. Purities of the unlisted nonpermethylated model peptides were >99%; purities of the unlisted permethylated compounds were >90%. Side chains of the unlisted compounds were unmodified by permethylation treatment.

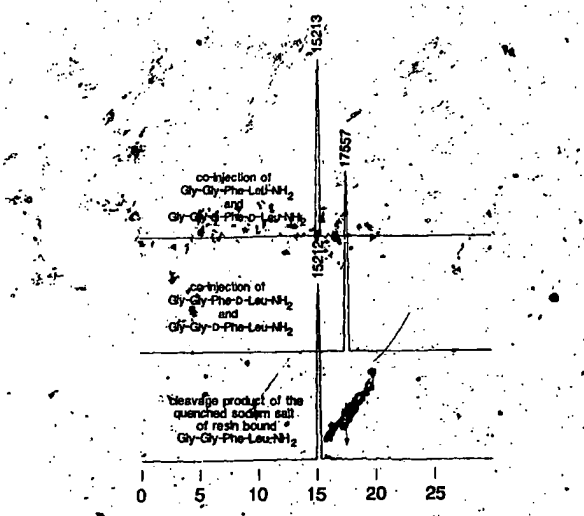


FIG. 3. RP-HPLC analysis of the four stereoisomers of GGFL-NH_2 after the formation and quenching of their amide anions by NaH. Arrow represents the retention time of the D/L and L/D stereoisomers.

cysteine and tryptophan were excluded). Each of the six permethylated sublibraries (represented as pm[*OXXXXX*], pm[*XOXXXX*], pm[*XXOXXX*], pm[*XXXOXX*], pm[*XXXXOX*], and pm[*XXXXXO*]) contained 37,791,360 (20×18^5) permethylated compounds in approximately equimolar amounts.

Antimicrobial Activity. Each of the 120 permethylated mixtures was assayed at an initial concentration of 2.5 mg/ml for its ability to inhibit the growth of *S. aureus*. A number of permethylated mixtures from each of the six libraries inhibited *S. aureus* growth (Fig. 4). None of the mixtures showed significant hemolytic activity or *in vitro* toxicity as determined by an MTT assay using McCoy cells ATCC 1696-CRL [MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] (25). The amino acids chosen at each of the defined positions from the most active permethylated mixtures were used to generate a series of individual permethylated compounds. Peptides were thus synthesized and permethylated representing all combinations of the amino acids chosen for the first position (W, F, Y, H, I, L), the second position (F), the third position (W, I, F), the fourth position (W, F), the fifth position (F, H), and the sixth position (F, H). The resultant 144 individual, crude permethylated compounds were screened for their antimicrobial activity. Forty-one compounds had IC_{50} values <50 μg/ml, whereas 51 compounds had IC_{50} values ≥250 μg/ml against *S. aureus*. Although the permethylated forms of FFIFFF-NH_2, FFFFFF-NH_2, and LFIFFF-NH_2 exhibited the greatest activity of the series (Table 2), the quaternary trimethyl ammonium salt of nonpermethylated FFFFFF-NH_2 showed no activity (IC_{50} >250 μg/ml). Similar activities were found against MRSA and *S. sanguis*, whereas none of the permethylated compounds or mixtures exhibited activity against *E. coli* or *C. albicans* (IC_{50} values >600 μg/ml) or showed significant toxicity as evidenced by lysis of red blood cells (<2% hemolysis at 100 μg/ml). These compounds showed activities similar to a range of previously described peptides made up of L-amino acids (1). However, in contrast to the L-amino acid peptides, these compounds appear completely stable to proteolytic enzymes. Mass spectral and RP-HPLC analysis showed that, in some cases, incomplete methylation had occurred. For the permethylated form of FFFFFF-NH_2, similar antimicrobial activity was found after purification, suggesting that incompletely methylated byproducts also had antimicrobial activity.

The potent antimicrobial activity of the permethylated hexaphenylalanine prompted a study to determine the length at which a permethylated polyphenylalanine has the greatest antimicrobial activity. The set of compounds prepared varied in length from one to eight residues. The antimicrobial activities of these compounds against the two Gram-positive bacteria, *S. aureus* and *S. sanguis*, are shown in Table 2. Significant activities were found for those permethylated compounds having a length of at least five residues, with a optimal length of seven residues. In addition, the seven- and eight-residue permethylated polyphenylalanine compounds exhibited similar activities against MRSA.

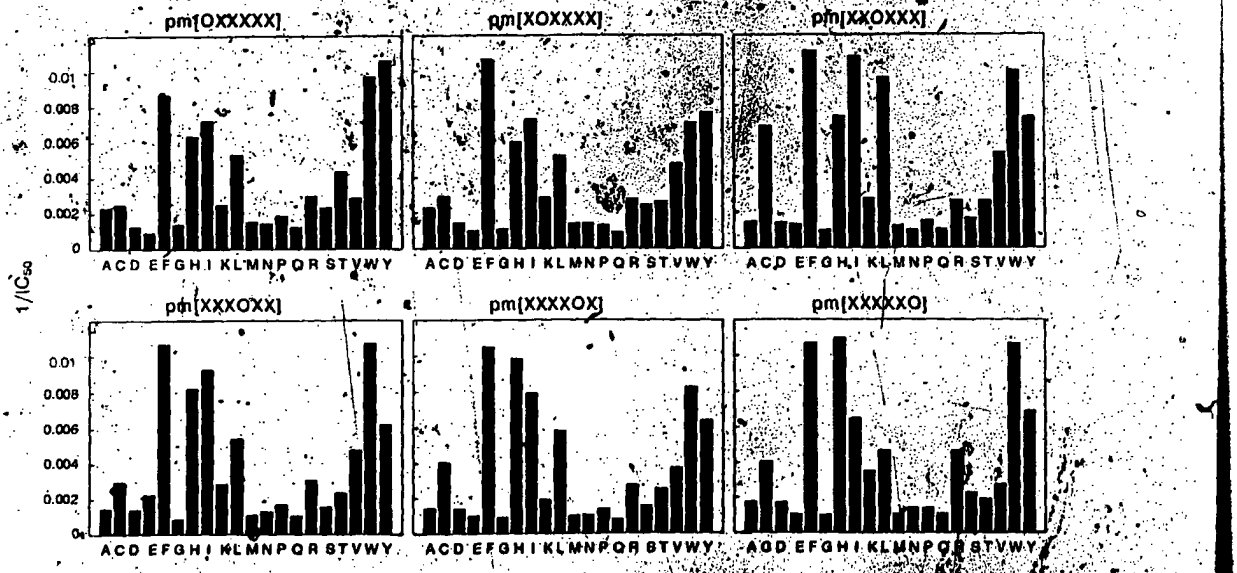


FIG. 4. Antimicrobial activity of the permethylated (pm-) PS-SPCL against *S. aureus*. Each individual bar represents the inverse of the IC_{50} of a permethylated peptide mixture defined in the "*O*" position with one of the 20 naturally occurring amino acids.

Table 2. Antimicrobial and hemolytic activities of permethylated peptides derived from a permethylated SPCL

Parent sequence	*S. aureus* IC_{50}, μg/ml	*S. aureus* MIC, μg/ml	*S. sanguis* IC_{50}, μg/ml	*S. sanguis* MIC, μg/ml
LFIFFF-NH_2	6	11–15	1.7	3–5
FFIFFF-NH_2	6	11–15	14	20–40
FFFFFF-NH_2	9	11–15	9	15–20
LFFFFF-NH_2	10	21–31	14	20–40
F-NH_2	>500	>500	>500	>500
FF-NH_2	>500	>500	>500	>500
FFF-NH_2	288	>500	446	>500
FFFF-NH_2	116	250–500	85	125–250
FFFFF-NH_2	19	21–31	20	25–30
FFFFFF-NH_2	7	11–15	5	7–8
FFFFFFF-NH_2	2.5	3–4	2.3	3–4
FFFFFFFF-NH_2	5	6–8	7	8–9

MIC, lowest concentration at which no growth is detected after 21-hr incubation.

DISCUSSION

Biologically active peptides can now be readily identified through the use of SPCLs and PS-SPCLs (1, 2, 11, 12). Peptides, however, have limitations as potential therapeutics; this is due to their lack of oral activity, rapid breakdown by proteolytic enzymes, rapid clearance from circulation, and typical inability to pass through the blood–brain barrier to effect central nervous system activity. The generation of libraries consisting of nonpeptidic compounds can be expected to circumvent a number of these limitations. We are pursuing a library approach in which existing and/or readily accessible peptide or other chemical diversities are chemically transformed to yield libraries (termed here chemically transformed libraries) having more desirable physical and chemical properties.

For the chemical transformation of peptide libraries to be of practical use, there are two requirements. (*i*) One must begin with a well-defined peptide library, and (*ii*) one must have access to chemical reagents that can effectively alter chemical moieties in a known manner, while leaving either all of the compound mixture on the resin or alternatively removing all of the mixture from the resin. The solid-phase synthesis of individual peptides (13, 18) and peptide libraries (1, 2) satisfies the first requirement because the preparation of libraries on solid supports can be done with a high degree of confidence and exactitude by recently described approaches (1, 2, 5, 12). Such libraries can then be chemically altered before their resin cleavage. The integrity of the peptide library used in the present study has been well demonstrated in earlier work (2, 12), and this library can be used for chemical transformation with confidence. The second requirement is satisfied here by the successful generation of chemically transformed peptide libraries. Initial studies with individual peptides indicate that peptides can be fully permethylated while bound to the resin support used in their synthesis. The average yield was >85% and was achieved independently of the sequence being permethylated.

Peptoid libraries (3), which are derived from the step-wise synthesis of amide-functionalized polyglycines, consist of compounds having a number of physical–chemical properties similar to the permethylated peptides described here (resistance to enzymes, favorable aqueous/lipid partitioning, etc.). We believe the use and transformation of existing resin-bound peptide libraries offer significant advantages, due to their general availability and to the familiarity of the procedures for the synthesis of peptides (13, 18) and peptide libraries (1, 2, 5, 12).

Thus, the transformation described here is an example of a more general approach to produce large diversities from readily accessible existing peptide libraries. We believe that the chemical transformation of a wide range of existing and future combinatorial libraries, as exemplified by the permethylated library presented here, offers a straightforward and rapid means to increase the available chemical diversity for use in basic research and drug discovery.

We thank Ema Takahashi for her expert technical assistance, Eileen Silva for help in document preparation. This work was funded by Houghten Pharmaceuticals, Inc., San Diego.

1. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature (London)* 354, 84–86.
2. Pinilla, C., Appel, J. R., Blanc, P. & Houghten, R. A. (1992) *Biotechniques* 13, 901–905.
3. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E. & Bartlett, P. A. (1992) *Proc. Natl. Acad. Sci. USA* 89, 9367–9371.
4. Geysen, H. M., Rodda, S. J. & Mason, T. J. (1986) *Mol. Immunol.* 23, 709–715.
5. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. (1991) *Nature (London)* 354, 82–84.
6. Eichler, J. & Houghten, R. A. (1993) *Biochemistry* 32, 11035–11041.
7. Scott, J. K. & Craig, L. (1994) *Curr. Opin. Biotechnol.* 5, 40–48.
8. DeWitt, S. H., Kiely, J. S., Stankovic, C. J., Schroeder, M. C., Cody, D. M. R. & Pavia, M. R. (1993) *Proc. Natl. Acad. Sci. USA* 90, 6909–6913.
9. Chen, C., Ahlberg-Randall, L. A., Miller, R. B., Jones, A. D. & Kurth, M. J. (1994) *J. Am. Chem. Soc.* 116, 2661–2662.
10. Bunin, B. A., Plunkett, M. J. & Ellman, J. A. (1994) *Proc. Natl. Acad. Sci. USA* 91, 4708–4712.
11. Dooley, C. T., Chung, N. N., Schiller, P. W. & Houghten, R. A. (1993) *Proc. Natl. Acad. Sci. USA* 90, 10811–10815.
12. Dooley, C. T. & Houghten, R. A. (1993) *Life Sci.* 52, 1509–1517.
13. Houghten, R. A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 5131–5135.
14. Blondelle, S. E. & Houghten, R. A. (1992) *Biochemistry* 31, 12688–12694.
15. Blondelle, S. E., Simpkins, L. R., Pérez-Payá, E. & Houghten, R. A. (1993) *Biochim. Biophys. Acta* 1202, 331–336.
16. Hakomori, S. I. (1964) *J. Biochem. (Tokyo)* 55, 205.
17. Challis, B. C. & Challis, J. A. (1970) in *The Chemistry of Amides*, ed. Zabicky, J. (Wiley, New York), pp. 731–857.
18. Merrifield, R. B. (1963) *J. Am. Chem. Soc.* 85, 2149–2154.
19. Ibrahim, H. H. & Lubell, W. D. (1993) *J. Org. Chem.* 58, 6438–6441.
20. Lawson, W. B., Gross, E., Foltz, C. M. & Witkz, B. (1962) *J. Am. Chem. Soc.* 84, 1715–1718.
21. Houghten, R. A., Bray, M. K., De Graw, S. T. & Kirby, C. J. (1986) *Int. J. Pept. Protein Res.* 27, 673–678.
22. Russell, G. A. & Weiner, S. A. (1966) *J. Am. Chem. Soc.* 31, 248–251.
23. Tam, J. P., Heath, W. F. & Merrifield, R. B. (1983) *J. Am. Chem. Soc.* 105, 6442–6455.
24. Piszkiewicz, D., Landon, M. & Smith, E. L. (1970) *Biochem. Biophys. Res. Commun.* 40, 1173–1177.
25. Mosmann, T. (1983) *J. Immunol. Methods* 65, 55–63.
26. Kornreich, W., Anderson, H., Porter, J., Vale, W. & Rivier, J. (1985) *Int. J. Pept. Protein Res.* 25, 414–420.